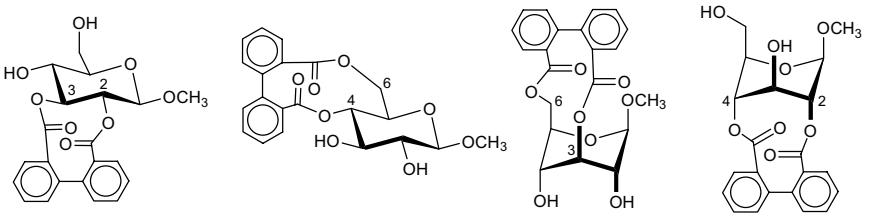
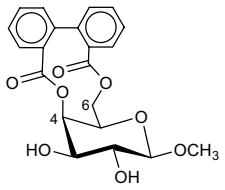
Chapter 8

Molecular Modelling of Saccharides



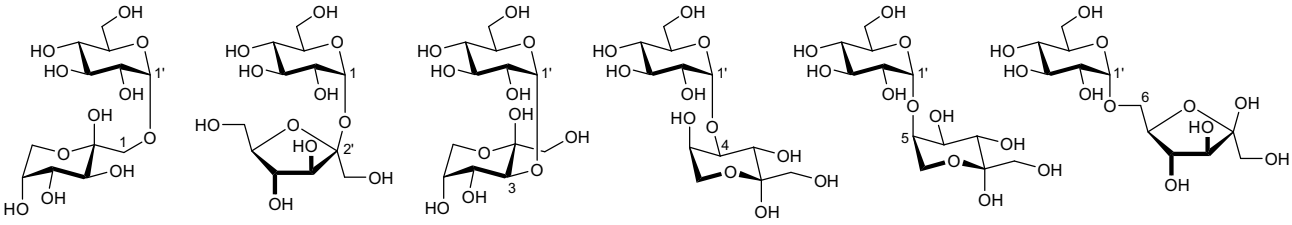
methyl *O*-diphenoyl-β-D-glucosides



methyl 4,6-*O*-diphenoyl-β-D-galactoside

Atropdiastereoisomers of Ellagitannin Model Compounds: Configuration, Conformation, and Relative Stability of D-Glucose Diphenoyl Derivatives

S. Immel, K. Khanbabaee,
Tetrahedron: Asymmetry **2000**, *11*, 2495-2507.



β-*p*-Trehalulose
α-D-Glc*p*-(1→1)-β-D-Fru*p*

Sucrose
β-D-Fru*f*-(2↔1)-α-D-Glu*p*

β-*p*-Turanose
α-D-Glc*p*-(1→3)-β-D-Fru*p*

β-*p*-Maltulose
α-D-Glc*p*-(1→4)-β-D-Fru*p*

β-*p*-Leucrose
α-D-Glc*p*-(1→5)-β-D-Fru*p*

β-*f*-Palatinose
α-D-Glc*p*-(1→6)-β-D-Fru*f*

Metabolism of Sucrose and Its Five Linkage-isomeric α-D-Glucosyl-D-fructoses by *Klebsiella pneumoniae*

J. Thompson, S. A. Robrish, S. Immel, F. W. Lichtenthaler, B. G. Hall, and A. Pikis,
J. Biol. Chem. **2001**, *276*, 37415-37425.

Metabolism of Sucrose and Its Five α-D-Glucosyl-D-fructose Isomers by *Fusobacterium mortiferum*

A. Pikis, S. Immel, S. A. Robrish, and J. Thompson,
Microbiology **2002**, *148*, 843-852.

Pergamon

Tetrahedron: *Asymmetry* 11 (2000) 2495–2507

TETRAHEDRON: *ASYMMETRY*

Atropdiastereoisomers of ellagitannin model compounds: configuration, conformation, and relative stability of D-glucose diphenoyl derivatives[1]

Stefan Immel[a,*] and Karamali Khanbabaee[b]

[a]*Institut für Organische Chemie, Technische Universität Darmstadt, Petersenstraße 22, D-64287 Darmstadt, Germany*
[b]*Fachbereich Chemie und Chemietechnik der Universität-GH Paderborn, Warburgerstraße 100, D-33098 Paderborn, Germany*

Received 18 April 2000; accepted 9 May 2000

Abstract

Conformational analysis reveals a remarkable rigidity of 2,3-, 4,6-, 3,6-, and 2,4-*O*-(*S*)- and (*R*)-diphenoyl (DP) bridged methyl β-D-glucosides, which were used as model compounds to evaluate the atropisomeric features of the natural ellagitannins, which possess at least one hexahydroxydiphenoyl (HHDP) moiety. The 2,3- and 4,6-*O*-(*S*)-DP bridged glucosides with $^4C_1$ pyranose geometries are thermodynamically more stable than their (*R*)-DP counterparts, whilst in the 3,6- and 2,4-*O*-linked series with $^1C_4$ glucopyranose geometries the (*R*)-DP configuration is preferred. The chiral scaffold of glucose exerts a strong atropdiastereoselective effect onto the diphenoyl units, which is mediated through 10- to 12-membered rings via ester linkages. The calculated results not only explain the observed (*S*)-diastereoselectivity of di-esterification reactions of suitably protected racemic hexaoxydiphenic acids with 4,6-unsubstituted D-glucopyranose derivatives, but also correlate the observed configuration of axially chiral HHDP-moieties of natural ellagitannins with conformational parameters.

1. Introduction

Ellagitannins constitute members of a large class of polyphenolic natural products that can be obtained by extraction from higher plants.[2] These extracts are widely used in folk medicine, leather and wine industry, and the tannin-components exhibit a broad range of biological activities.[2] The common structural element of the ellagitannins is a hexahydroxydiphenoyl (HHDP) unit located at the 2,3-, 4,6-, 1,6-, 3,6-, and/or 2,4-positions of the D-glucopyranose core.

The axially chiral HHDP-units of the 2,3- and 4,6-*O*-HHDP ellagitannins exhibit almost invariably the (*S*)-configuration with only very few exceptions (Cercidinin A and B,[3] Cuspinin,[3] and Platycaryanin D[4]) amongst more than 500 structurally characterized compounds (Fig. 1).[2,5]

* Corresponding author. E-mail: lemmi@sugar.oc.chemie.tu-darmstadt.de

0957-4166/00/$ - see front matter
PII: S0957-4166(00)00179-8

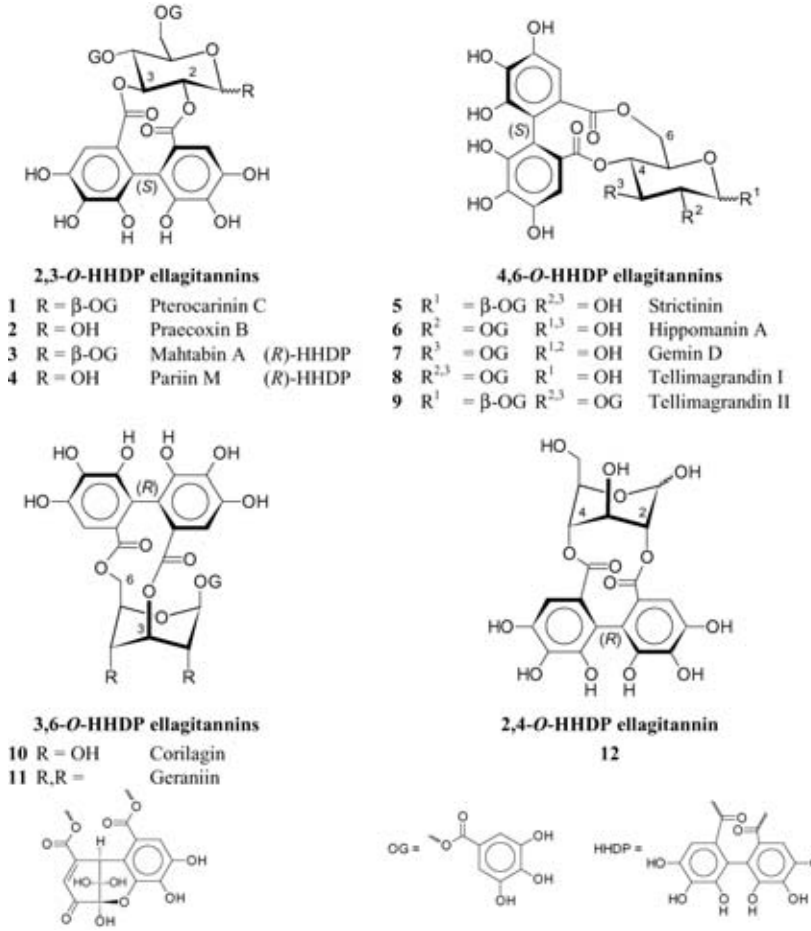


Figure 1. Chemical formulas of 2,3-, 4,6-, 3,6-, and 2,4-*O*-hexahydroxydiphenoyl (HHDP) glucosides (ellagitannins) with axial chirality of the HHDP moieties

Recently, it was shown that the postulated structures for the unusual ellagitannins Cercidinin A and B,[3] each possessing a 2,3-(*R*)-HHDP unit, are incorrect and must be revised.[6]

Prominent examples for the 2,3-*O*-(*S*)-HHDP ellagitannins are Pterocarinin C **1** and Praecoxin B **2**,[7] their non-natural, unusual (*R*)-HHDP counterparts Mahtabin A **3** and Pariin M **4** have become accessible in enantiomerically pure form through ring forming di-esterification of the *o*-nitrobenzyl 4,6-*O*-benzylidene-β-D-glucoside with racemic hexabenzyloxydiphenic acid.[6] Total syntheses of the natural 4,6-*O*-(*S*)-HHDP ellagitannins **5**–**7**[8,9] were achieved through unexpectedly highly atropdiastereoselective di-esterification reactions of racemic hexabenzyloxydiphenic acid with different 4,6-unsubstituted D-glucopyranose derivatives, exclusively leading to the corresponding 4,6-*O*-(*S*)-HHDP diastereomers. In contrast, in the same reactions the (*R*)-component of the racemic hexabenzyloxydiphenic acid gave rise to the formation of oligomers through the possible intermolecular competition pathway only.[8–10] On the other hand, the naturally occurring ellagitannins in which the HHDP-units are attached to the 3,6- (Corilagin **10**[11] and Geraniin **11**[12,13]) or 2,4-positions[14] **12** of glucopyranose exhibit (*R*)-HHDP configuration,[2] although the latter invariably undergo further biochemical transformations under physiological conditions.[2]

The highly atropdiastereoselective formation of the ellagitannins from their biochemical gallotannin precursors has been explained through conformational preferences of the galloyl residues, and strain in the transition states of the oxidative coupling reactions between galloyl esters.[2,12,15–17] However, despite the wide-spread occurrence of polyphenolic compounds of tannin class natural products, very little structural data on gallotannins or ellagitannins at atomic resolution is available through crystal structure analysis; the only examples available from the Cambridge Crystallographic Database[18] are Geraniin **11**[19] and methyl 4,6-*O*-benzylidene-2,3-*O*-(*S*)-hexamethoxydiphenoyl-α-D-glucopyranoside **13**.[20]

As chemical syntheses of the ellagitannins proceed either via oxidative coupling of galloyl residues in galloylated glucose substrates, or alternatively through di-esterification of methyl- or benzylether-protected enantiopure or racemic hexahydroxydiphenic acid with appropriately substituted D-glucosides,[2] stereocontrol of these reactions is of fundamental importance. Therefore, we have undertaken a molecular modeling study of some ellagitannin model compounds with the aim of obtaining structural models on an atomic level, and to explain the relative stabilities of the various (*R* and *S*)-HHDP diasteromers under equilibrium conditions.

2. Results and discussion

We chose the compounds **14**–**21** (Fig. 2) as a starting point for our study on the different types of ellagitannins. The hydroxyl groups of the HHDP- and galloyl-residues were omitted in order to reduce the number of local minima on the energy potentials surfaces of the ellagitannins originating from different OH-rotamers (→diphenoyl and benzoyl-substituents). This simplification also avoids conformational artifacts stabilized by strong intramolecular hydrogen bonds, as the geometry analyses were carried out for the isolated molecules only without the explicit incorporation of a solvent; for the same reason, the methyl β-D-glucosides were considered only. Starting geometries were generated for all types of 2,3-, 4,6-, 3,6-, and 2,4-type linkages, and both (*S*)- and (*R*)-diphenoyl (DP) atropdiastereoisomers, respectively. In the first class of compounds, the 4- and 6-OH groups of glucose were 'blocked' by benzoyl groups as mimics for the galloyl residues (vide supra).

In each case, the conformational space was explored by a mixed molecular dynamics (MD) and molecular mechanics (MM) approach, applying a full energy optimization to 5000 structures extracted along a 500 ps MD trajectory (for details, see Experimental). As was established through monitoring a number of molecular parameters along the MD runs, this methodology did not only produce glucose conformations other than $^4C_1$ or $^1C_4$, but it also generated a large number of conformers of the diphenoyl ring system. The atropstereochemical configuration of all diphenoyl units was retained during all MD simulations, and no (*S*)↔(*R*) transitions were recorded. The fully relaxed, global energy-minimum structures obtained were used in this report, and for each model compound some characteristic geometry parameters (cf. Fig. 2) are listed in Table 1.

2.1. 2,3-O-Diphenoyl glucosides *14* and *15*

The global energy-minimum structures of **14** and **15** are shown in Fig. 3, with **14** being about 7.4 kJ/mol more stable than **15** (cf. Table 1). The glucopyranose units adopt standard $^4C_1$ ring geometries as evidenced by their Cremer-Pople ring puckering parameters[21] (cf. Table 1), and the

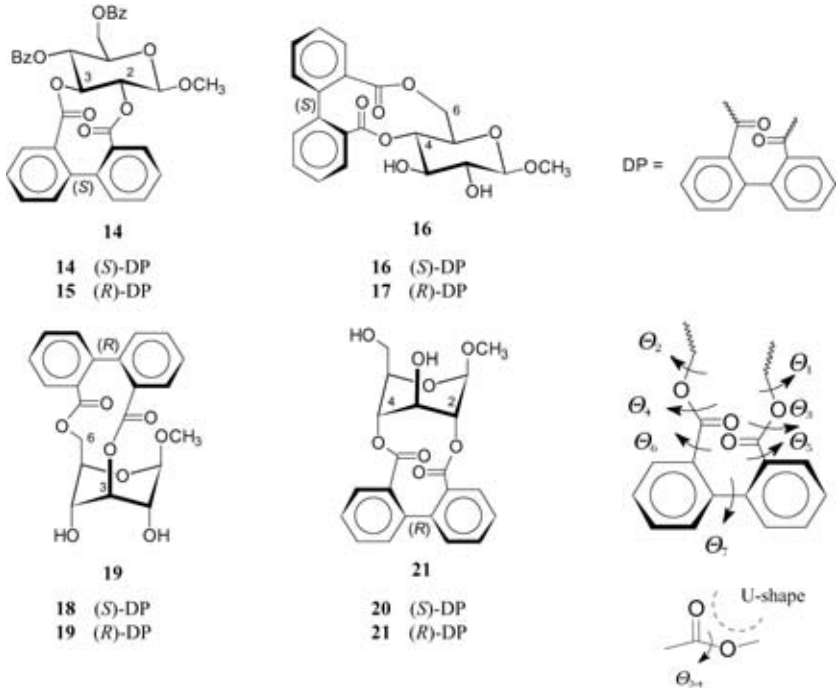


Figure 2. Methyl 2,3-, 4,6-, 3,6-, and 2,4-*O*-(*R*,*S*)-diphenoyl-β-D-glucosides **14**–**21** used as model structures for evaluating the conformational properties of ellagitannins (DP = diphenoyl). On the right, some ring torsion angles Θ_1–Θ_7 used in the geometry analysis are given. Θ_1 and Θ_2 define the mode of attachment of the diphenoyl-unit to the glucopyranose ring. The preferred U-shape of both ester groups is characterized by Θ_3 and Θ_4 with ideal values of approximately 0°; Θ_5 and Θ_6 denote the inclination of the carbonyl groups towards the phenyl rings (ideal values 0° and ±180° for conjugated π-systems). The atropisomeric diphenoyl units display torsion angles Θ_7 of opposite sign (*R*: negative, *S*: positive values); Θ_7 is approximately equivalent to the tilt angle between the phenyl rings

conformational preferences of the 10-membered diphenoyl ring system are largely determined by its rather rigid *trans*-type linkage to the pyranose scaffold. The two ester groups display a highly characteristic tendency to maintain an U-shape (Fig. 2), although the torsion angles Θ_3 and Θ_4 indicate about a 30° deviation from the ideal geometry for **14** (Table 1). In **15**, the glucose 2-OCO-ester group is forced by the (*R*)-DP residue into a strained *trans*-type conformation with $\Theta_3 \approx 140°$. In both compounds **14** and **15**, the carbonyl groups are inclined by about 40–50° (Θ_5 and Θ_6) relative towards the planes of phenyl ring π-systems. The phenyl rings of the DP-units are tilted towards each other by about 50–60° (torsion angle Θ_7 in Table 1, positive values of Θ_7 indicating (*S*)-DP units and negative (*R*)-DP configurations). Most notably, the ester groups tend to adopt an antiparallel arrangement of their C=O dipoles (angles φ of 150–165°), whereas the tilt angles τ indicate an almost parallel (stacked) alignment of the planes formed by the C-COO-atoms of each ester fragment. These geometry parameters clearly reveal the torsion angle Θ_3 as the main reason for the lower stability of **15** as compared to **14**.

This notion is further substantiated through color-coded projection of the force-field derived split-terms of strain energy originating from angle- and torsion-bending onto the ball-and-stick models of **14** and **15**. In Fig. 4, blue colors correspond to relaxed molecular fragments, whereas yellow to red colors indicate strain on distinct residues. In both cases, the type of ring-anellation

Table 1
Relative energies and selected geometry parameters calculated for the global energy-minimum structures of the ellagitannin model compounds **14**–**21**. For each type of 2,3-, 4,6-, 3,6-, and 2,4-*O*-diphenoyl bridged glucopyranose derivative, the (*S*)- and (*R*)-atropdiastereomers are listed; the *galacto*-configured compounds **23** and **24** lack naturally occurring counterparts amongst the ellagitannins, but were included for comparison (cf. text)

	compound	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**23**	**24**
	linkage type	2,3-(*S*)	2,3-(*R*)	4,6-(*S*)	4,6-(*R*)	3,6-(*S*)	3,6-(*R*)	2,4-(*S*)	2,4-(*R*)	4,6-(*S*)[a]	4,6-(*R*)[a]
ΔH_{calc} [kJ/mol]	absolute	-	-	-	-	-	-	-	-	-	-
	relative	1245.4	1238.0	1158.7	1153.2	1145.3	1172.4	1125.3	1130.9	1150.9	1159.5
		0.0	7.4	0.0	5.5	27.1	0.0	5.6	0.0	8.6	0.0
	conformation	$^4C_1$	$^4C_1$	$^4C_1$	$^4C_1$	$^1C_4$	$^1C_4$	$^1C_4$	$^1C_4$	$^4C_1$	$^4C_1$
pyranose Cremer-Pople parameters[21]	Q [Å]	0.545	0.615	0.556	0.557	0.490	0.528	0.543	0.543	0.577	0.545
	θ [°]	11.6	8.4	11.5	7.3	176.2	172.6	173.6	172.6	3.8	5.4
	ϕ [°]	37.8	243.8	341.2	285.1	208.0	31.2	43.3	317.8	288.1	21.9
torsion angles[b] [°]	Θ_1 (C_{pyr}-C''$_{pyr}$-O-C)	-81.6	-84.3	-130.9	-162.8	73.2	75.2	-64.1	-68.9	164.5	128.8
	Θ_2 (C'$_{pyr}$-C''$_{pyr}$-O-C)	-82.1	48.2	-111.8	91.5	118.3	125.2	58.1	63.7	116.8	115.3
	Θ_3 (C_{pyr}-O-C=O)	-32.7	136.9	-28.9	-116.0	53.9	0.1	-20.5	-166.0	130.5	26.4
	Θ_4 (C''$_{pyr}$-O-C=O)	-33.0	34.1	-30.5	18.8	28.6	40.5	-176.1	23.5	177.1	29.5
	Θ_5 (O=C-C_{dp}-C_{dp})	42.4	-44.6	49.2	-107.0	138.3	-43.5	63.5	-61.1	124.6	-47.6
	Θ_6 (O=C-C'$_{dp}$-C'$_{dp}$)	42.9	-50.9	36.9	-38.3	141.7	-33.6	49.9	-60.6	78.4	-38.9
	Θ_7(C_{dp}-C_{dp}-C'$_{dp}$-C'$_{dp}$)	48.9	-61.9	56.2	-74.6	90.8	-58.3	63.9	-70.1	114.6	-54.5
	ω (O_5-C_5-C_6-O_6)	43.7[c]	71.0[c]	-80.9	-170.7	-99.0	175.3	173.0[c]	167.6[c]	168.8	-173.9
angle [°]	φ (C=O / C=O)[d]	165.9	153.3	150.6	120.7	167.1	154.1	138.9	138.2	130.4	151.4
tilt angle [°]	τ (COO / COO)[e]	175.6	176.2	168.5	132.4	88.0	160.6	162.8	167.3	134.4	168.9

a) D-*galacto*-configuration. — b) C_{pyr} and C'$_{pyr}$ denote the pyranose positions linked to the diphenoyl moiety, i.e. C_{pyr} / C'$_{pyr}$ = C_2 / C_3 for **14** and **15**; C_4 / C_6 for **16, 17, 23,** and **24**; C_3 / C_6 for **18** and **19**; C_2 / C_4 for **20** and **21**; C_{dp} and C'$_{dp}$ refer to the diphenoyl unit. — c) torsion angle ω not part of the diphenoyl ring system in **14**, **15**, **20** and **21**. — d) angle between the bond vectors of the ester carbonyl groups. — e) tilt angle between the two planes defined by the ester groups (atoms C-COO).

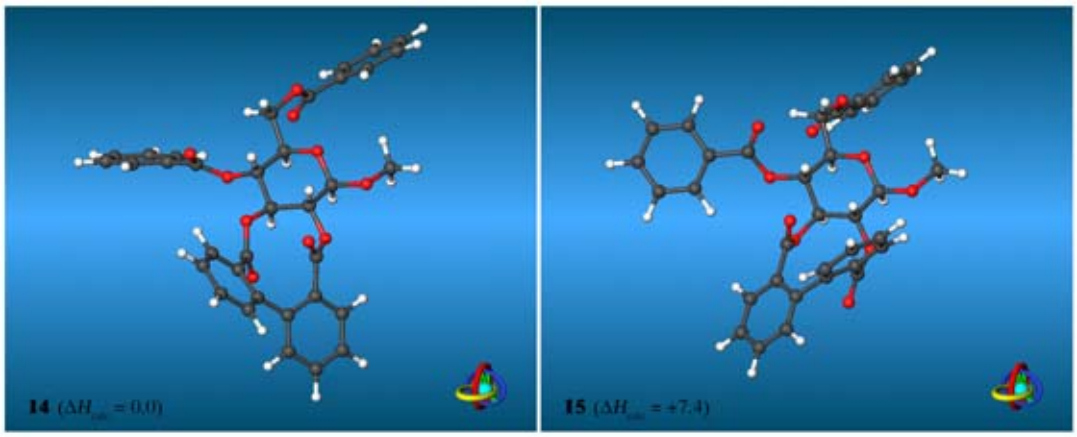


Figure 3. Ball-and-stick models of the global energy-minimum structures of the 2,3-*O*-(*S*)-DP **14** and (*R*)-DP **15** glucosides; calculated relative energies are given in kJ/mol

exerts strain on the position C-4 of the glucopyranose. In addition, **15** clearly displays through red colors internal strain centered around the 2-OCO ester. The rather rigid glucose unit serves as chiral strait-jacket, to which only a (*S*)-DP unit can be attached in a 'relaxed' conformation to the 2-O- and 3-O-groups, the atropstereochemical induction being mediated through the stiff ester groups in the attached ring.

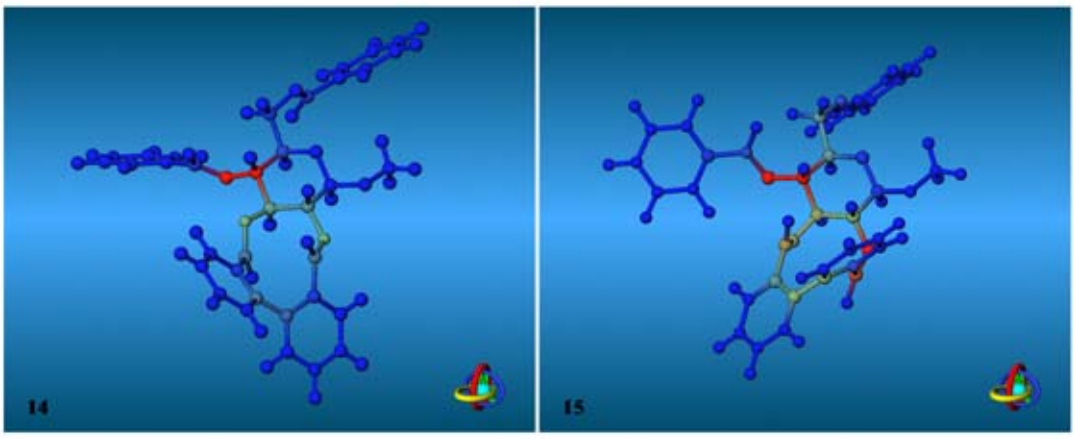
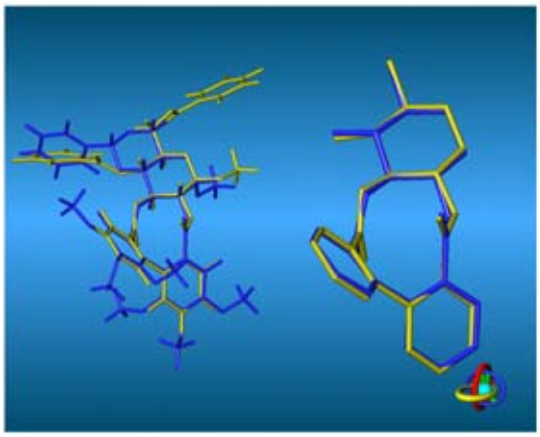
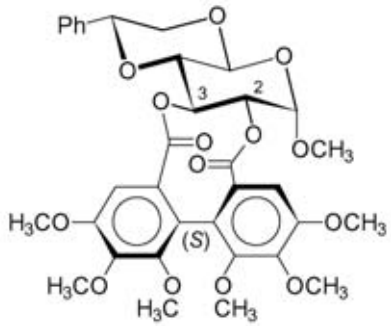


Figure 4. Color-coded projection of the force-field derived sum of split terms of angle- and torsion-bending strain energy onto ball-and-stick models of 2,3-*O*-(*S*)-DP **14** (left) and (*R*)-DP **15** (right) glucosides; blue colors indicate relaxed molecular parts, and yellow/green to red colors designate strained fragments (strain energies 0–5 kJ/mol); the mode of viewing corresponds to Fig. 3. The rigidity of both the pyranose and DP-ring systems, as well as the stiffness of the ester groups lead to a high energy conformation of the glucose 2-O-ester linkage in **15**

Some indications on the relevance of the computer-generated geometries are derived from comparison of **14** with the conformation of methyl 4,6-*O*-benzylidene-2,3-*O*-(*S*)-hexamethoxydiphenoyl-α-D-glucopyranoside **13** obtained from crystal structure analysis.[20] Superimposition of the common molecular fragments of **14** and **13** (Fig. 5) displays a high degree of correlation between the theoretically predicted and experimentally observed structures: in particular the ring linkage, the alignment of the ester groups, as well as the relative tilt of the phenyl rings are predicted accurately.





23

Figure 5. Superimposition of the computer-generated geometry of **14** (yellow model) with the solid-state conformation of methyl 4,6-*O*-benzylidene-2,3-*O*-(*S*)-hexamethoxydiphenoyl-α-D-glucopyranoside[20] (**13**, blue model). On the left, both structures are displayed entirely, whereas on the right only the common molecular fragments used for 3D-fitting are shown in enlarged form; the best-fit yielded a root mean square (RMS) deviation of $\sigma = 0.08$ Å for these atomic positions

2.2. 4,6-O-Diphenoyl glucosides **16** and **17**

In the 4,6-*O*-DP bridged glucosides **16** and **17** (Fig. 6), the same basic conformational effects as in the 2,3-*O*-DP series (**14** and **15**) are operative: the rather stiff U-shape of both ester groups, the preferred antiparallel C=O-dipole–dipole alignment, the conjugation of carbonyl and phenyl ring π-systems, as well as the tilt between the phenyl rings. In addition, the glucose 6-CH_2O-group is less flexible than it appears: of the three principal staggered conformations of the 6-C–O-linkage, the *trans–gauche* (*tg*)[22] form is commonly the least preferred one, as it is destabilized by 1,3-diaxial like repulsions between O-4 and O-6 (Scheme 1).[22,23] Although the *gauche–gauche* (*gg*) and *gauche–trans* (*gt*) forms[22] are usually equivalent in energy in non-cyclic glucose derivatives,[23] the latter becomes inaccessible in the 11-membered ring systems of the 4,6-*O*-DP glucosides **16** and **17** for sterical reasons. The *gg*-form[22] is realized in **16** ($\omega \approx -80°$, cf. Table 1), still allowing almost relaxed U-shaped ester linkages (Θ_3 and $\Theta_4 \approx -30°$). As discussed above, a value of $\Theta_3 \approx -116°$ indicates a highly bent 6-OCO-ester linkage in **17**, and its strain is only partly relaxed through sacrificing the 6-O *gg*-form to a less preferred *tg*-geometry ($\omega \approx -170°$ in **17**). However, superimposition of multiple low-energy conformers for both **16** and **17** (Fig. 7) indicates a lower flexibility of **16** as all ring fragments reside in a 'relaxed' state, whereas the strain in **17** (vide supra) allows for multiple conformations of the 6-C-OCO-fragments with different orientations of the ester carbonyl group. The data listed in Table 1 also indicates a less favorable alignment of the ester groups in **17** as compared to **16** (angles of $\varphi \approx 120°$ vs. 150° and tilts $\tau \approx 130°$ vs. 170° in **17** and **16**).

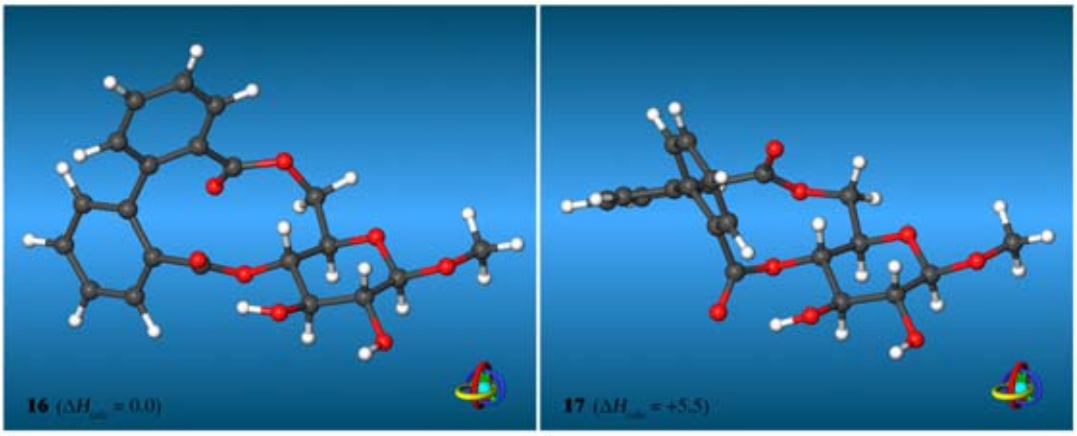
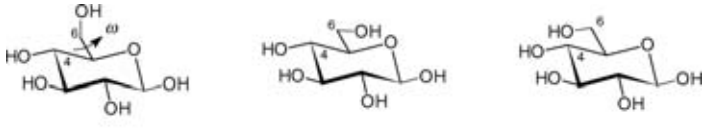


Figure 6. Ball-and-stick models of the global energy-minimum structures of the 4,6-*O*-(*S*)-DP **16** and (*R*)-DP **17** glucosides; relative energies are given in kJ/mol



Scheme 1.

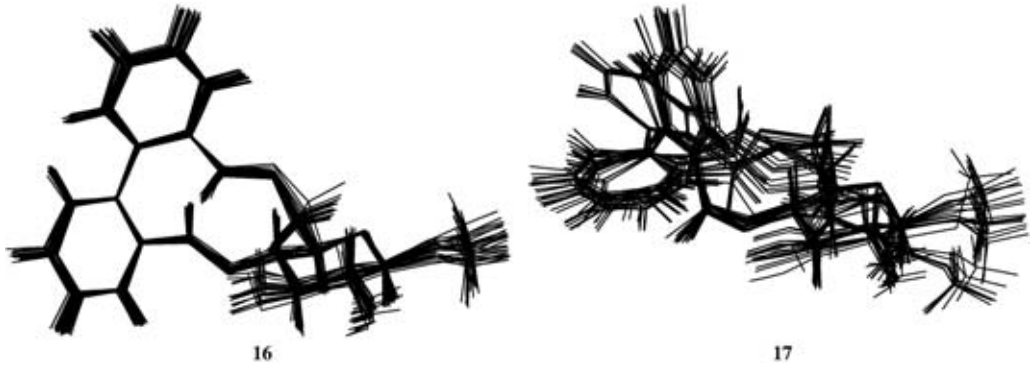
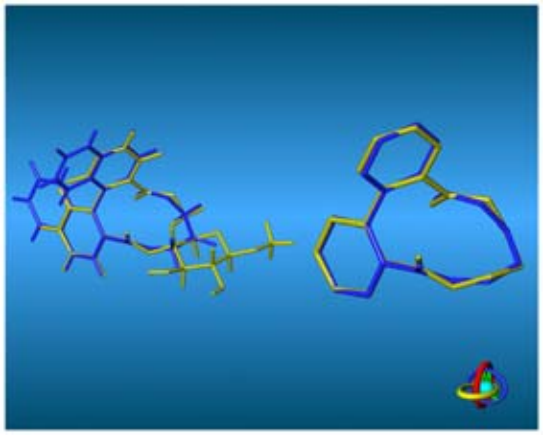
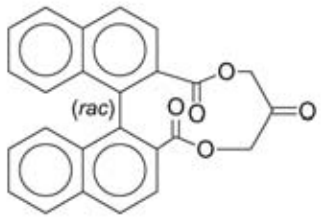


Figure 7. Superimposition of the 25 most stable geometries found within an energy range of approx. 0–10 kJ/mol above the global energy-minima for the atropdiastereomeric 4,6-*O*-DP glucoside model compounds **16** and **17**, respectively; the viewpoint relative to the glucopyranose ring is identical in both cases. The more stable (*S*)-DP compound **16** is remarkably rigid with rather stiff ring conformations. (*R*)-**17** displays some flexibility, and particularly the 6-CH_2-OCO ester carbonyl group adopts different orientations with almost equal energies (cf. text)

In Fig. 8, the calculated geometry of **16** is compared with the corresponding (*S*)-binaphthalene conformation in the crystal structure of racemic **22**:[24] although the 11-membered ring in the latter compound includes a Csp^2 carbonyl bridge head versus a 5-Csp^3 carbon in D-glucose, the 3D-fit displays very similar ring conformations in both compounds. In total, the geometry parameters obtained are consistent with preferred 4,6-*O*-(*S*)-DP D-glucose type linkages over (*R*)-DP geometries, and thus agree well with the absence of natural ellagitannins of the latter type.



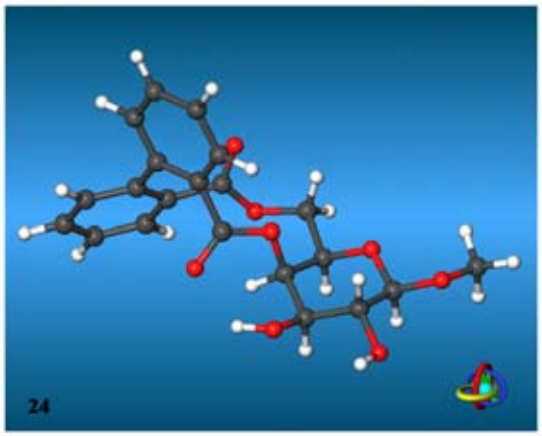


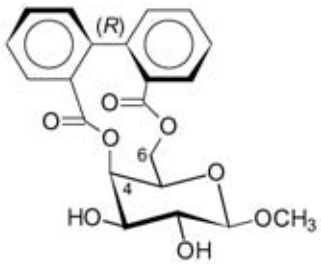
22

Figure 8. Superimposition of the 3D-structures of methyl 4,6-*O*-(*S*)-diphenoyl-β-D-glucoside (**16**) with the corresponding (*S*)-binaphthalene geometry extracted from the solid-state structure of racemic **22**,[24] the mode of viewing corresponds to Fig. 5. Despite the chemical differences of both compounds, the 11-membered rings display similar conformations with RMS deviations in the atomic positions of $\sigma \approx 0.1$ Å only

It is noteworthy, that the atropstereochemical induction on the DP-units strongly depends on the type of linkage and the configuration of the carbohydrate template: comparison of the 4,6-*O*-(*S*)-DP and 4,6-*O*-(*R*)-DP D-galactopyranosides **23** and **24**[25] with their D-glucose analogs **16** and **17** reveals opposite effects on the chirality of the DP-residues. The (*R*)-configuration of **24** is

about 8.6 kJ/mol more stable than the corresponding (*S*)-geometry (cf. Table 1 and Fig. 9). Due to the axial configuration at C-4 in D-galactose, the conformational preference of the 6-CH_2O-group is shifted from the *gg*- to the *tg*-form.[22]





23 (*S*)-DP
24 (*R*)-DP

Figure 9. Global energy-minimum structure of 4,6-*O*-(*R*)-DP galactopyranoside **24** (left)

*2.3. 3,6- and 2,4-*O*-Diphenoyl glucosides **18–21***

In both series of 3,6-*O*-DP (**18**, **19**) and 2,4-*O*-DP (**20**, **21**) bridged methyl β-D-glucosides, the glucose units are forced into $^1C_4$ chair geometries to affect ring closure (cf. Cremer–Pople ring puckering parameters[21] in Table 1); a behavior that is well-documented in the chemistry of 3,6-anhydro glucose derivatives.[26] Thus, the linkage positions as well as all other ring substituents adopt axial orientations on the pyranose ring, and in both 3,6-*O*-DP and 2,4-*O*-DP glucosides the (*R*)-DP configuration **19** and **21** is energetically preferred over the (*S*)-DP analogs **18** and **20** (cf. Table 1, and Figs. 10 and 11).

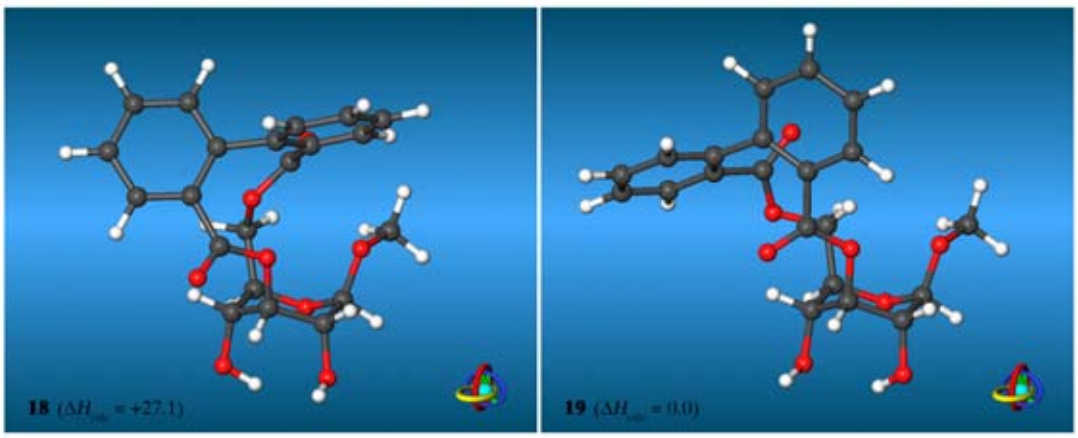


Figure 10. Global energy-minimum structures of the 3,6-*O*-(*S*)-DP **18** and (*R*)-DP **19** glucosides; relative energies are given in kJ/mol

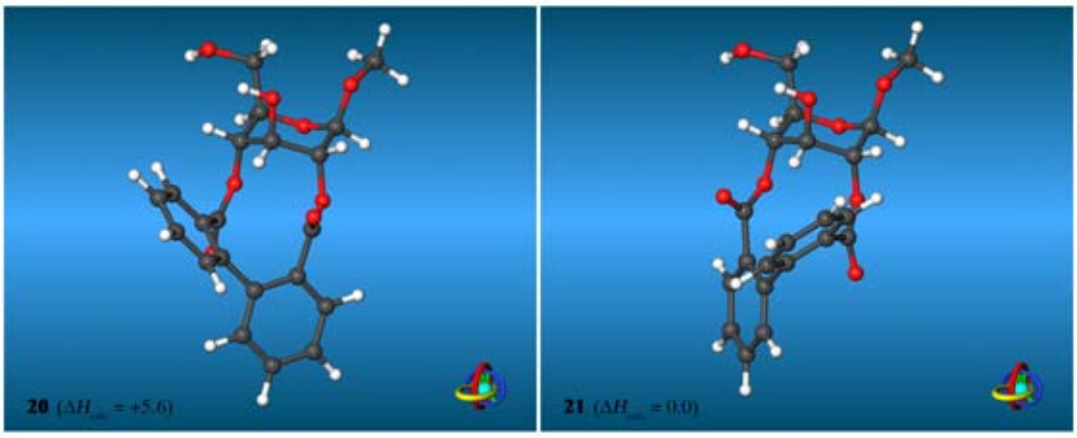
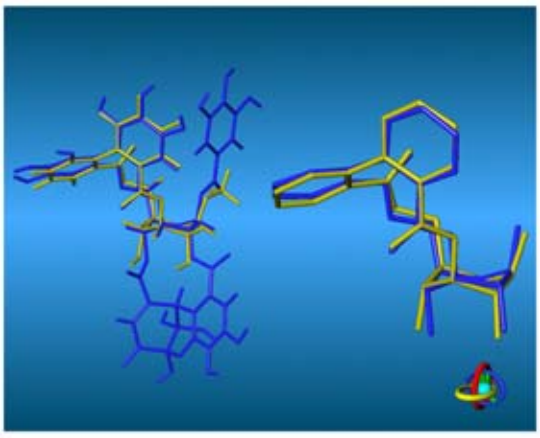
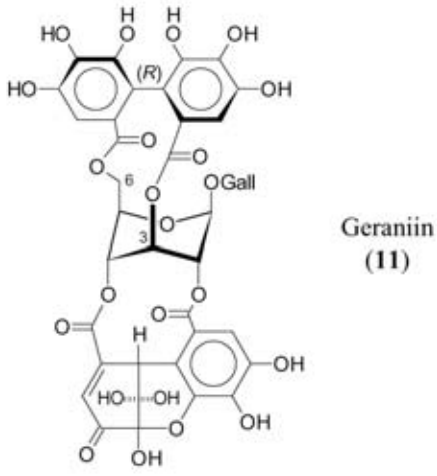


Figure 11. Global energy-minimum structures of the 2,4-*O*-(*S*)-DP **20** and (*R*)-DP **21** glucosides

As discussed above, it is the high tendency of the ester groups to maintain their characteristic U-shape, that transmits the chiral information of the carbohydrate backbone into the DP-residue, and thus induces their atropstereochemical configuration. Comparison of the calculated structure of **19** with the solid-state conformation of the natural ellagitannin Geraniin **11**[19] (Fig. 12) reveals very similar over all-shapes for both 3,6-*O*-(*R*)-DP ring systems, although the puckering of the pyranose ring is calculated to be smaller than found in the crystal structure of **11**.[19] However, the ring flattening effect observed in **19** is counterbalanced in **11** through an additional rigid linkage between the axial 2- and 4-OH groups of the D-glucoside core.





Geraniin
(11)

Figure 12. Superimposition of the energy-minimum structure **19** (yellow model) with the solid-state geometry of Geraniin **11**[19] (blue colors); the mode of viewing is analog to Figs. 5 and 8 (RMS deviation of $\sigma = 0.11$ Å for all equivalent atomic positions)

It is noteworthy, that in neither the 2,4-*O*-(*S*)- or (*R*)-DP glucosides **20** and **21** both ester groups are capable to maintain their typical U-shape simultaneously. The internal strain inherent to these compounds may be considered one reason why their ellagitannin analogs are not stable under physiological conditions and undergo very characteristic transformations into ring systems as realized in **11**[2] (Figs. 1 and 12; in the solid-state geometry of **11** the ester U-shape is realized in both the 2-OCO and 4-OCO linkages). However, the structure of Geraniin **11**, and the stereochemical course of this process indicates, that 2,4-*O*-(*R*)-HHDP units are the precursors rather than (*S*)-HHDP diastereomers.[2] The observation that these (*R*)-HHDP intermediates have to play an important role along the biochemical pathways of some ellagitannins correlates well with the observed stability of 2,4-*O*-(*R*)-DP glucosides over the corresponding (*S*)-DP geometries.

3. Conclusion

Without exception, the calculated relative thermodynamic stabilities of the atropdiastereomeric methyl 2,3-, 4,6-, 3,6-, and 2,4-*O*-DP-β-D-glucosides **14**–**21** correlate with the configuration of the corresponding naturally occurring HHDP-bridged ellagitannins,[2] i.e. the 2,3- and 4,6-*O*-(*S*)-DP, as well as the 3,6- and 2,4-*O*-(*R*)-DP glucosides are more stable than their diastereomers. The molecular modeling data and the calculated energy differences are consistent with limited flexibility of bis-equatorial linked DP-residues in 2,3- and 4,6-positions of D-glucose ($^4C_1$ pyranose conformation), the 10-membered ring in the 2,3-*O*-DP glucosides being less flexible than the 11-membered 4,6-*O*-DP macrocycle with one additional CH_2-group. Even more rigid are the bis-axial linked 3,6-*O*- and 2,4-*O*-DP glucosides with $^1C_4$ glucose *anti*-chair geometries, the latter bearing even so much internal strain that they immediately undergo ensuing biochemical transformations under physiological conditions (vide supra).[2] Comparison of the calculated structures with the few solid-state structures available for **11**, **13**, and **22** attests to the consistency of the methodology applied. The strait-jacket of the carbohydrate scaffold exerts a strong chiral induction onto the attached diphenoyl-residues in all types of linkages discussed within this study, the major effect in chirality transduction emerging from the tendency of the ester groups to maintain their characteristic U-shape. Indeed, preliminary results indicate that the atropchiral induction is significantly weakened or even inverted[27] when replacing the -COO-ester groups with -CH_2-O-type linkages. As detailed for the non-natural methyl 4,6-*O*-DP galactosides **23** and **24**, the configuration of the DP-units is highly sensitive towards changes in the linkage type and/or geometry.

On the basis of the molecular geometries discussed here, not only molecular models for the various ellagitannins can be derived, but also the configuration of the HHDP-units found in the class of ellagitannins can be rationalized. Although biosynthesis of the ellagitannins proceeds via HHDP ring formation through oxidative coupling of galloyl residues, and thus the strain inherent to the transition states of ring closure is the decisive kinetic factor for the atropdiastereoselective formation of (*S*)- or (*R*)-HHDP ellagitannins, the transitions states may somewhat 'resemble' the final ellagitannin products. For this reason, the obvious build-up of torsional strain in one of the corresponding diastereomers each is apt to explain the very uniform patterns of HHDP-configurations found in the natural ellagitannins.

Oxidation of the ellagitannins at C-1 (→gluconolactons) and subsequent strain-induced pyranose ring opening[10] leads to open-chain tannins such as the Lagerstannin A, B, and C.[28] We hope to report on their geometries and conformational preferences in the near future.

4. Experimental

4.1. Crystal structures, fitting, and molecular graphics

The solid-state structures of Geraniin **11**,[19] **13**,[20] and **22**[24] were extracted from the Cambridge Crystallographic Database.[18] Hydrogen atoms not included in the structure determinations or disordered positions were positioned geometrically. 3D-fitting of molecular geometries was carried out by rigid-body translation and rotation, considering all equivalent non-hydrogen atoms with equal weights. All molecular graphics were generated using the MolArch+ program.[29]

4.2. Molecular dynamics (MD) and molecular mechanics (MM)

All MD and MM calculations were carried out using the PIMM91 force-field[30] without the explicit incorporation of solvent molecules. All starting geometries of compounds **14**–**21** were generated using the MolArch+ program,[29] different starting geometries for **16** and **17** were used to ensure the self-consistency of the search procedure for the global energy-minimum structures. For each compound, 500 ps MD trajectories were generated (time step $\Delta t = 0.5$ fs, $T = 300$ K), and molecular configurations were saved every 200 steps. Monitoring the glucopyranose geometries, as well as torsion angles along the DP-ring systems was used to ensure comprehensive coverage of the conformational space by the MD runs. For each compound **14**–**21**, a total of 5000 MD snapshot conformations were subjected to full MM energy minimization, and the global energy-minimum conformations found in each case were used in this report, and all geometry parameters (cf. Fig. 2) listed in Table 1 were calculated from this data set.[29] It is noteworthy, that the global energy-minimum structures were found multiple times in all cases, and that most of the low energy conformers within a range of 0–5 kJ/mol cluster into the same conformational family as the most stable geometry (for example, see Fig. 7).

References

1. Part 28 of the series Molecular Modelling of Saccharides. For Part 27, see: Immel, S.; Nakagawa, T.; Lindner, H.-J.; Lichtenthaler, F. W. *Chem. Eur. J.* **2000**, *6*, in press.
2. For a review, see: Quideau, S.; Feldmann, K. S. *Chem. Rev.* **1996**, *96*, 475–503.
3. Nonaka, G.-I.; Ishimatsu, M.; Ageta, M.; Nishioka, I. *Chem. Pharm. Bull.* **1989**, *37*, 50–53.
4. Tanaka, T.; Kirahara, S.; Nonaka, G.-I.; Nishioka, I. *Chem. Pharm. Bull.* **1993**, *41*, 1708–1716.
5. Feldmann, K. S.; Smith, R. S. *J. Org. Chem.* **1996**, *61*, 2606–2612.
6. Khanbabaee, K.; Lötzerich, K. *J. Org. Chem.* **1998**, *63*, 8723–8728.
7. Khanbabaee, K.; Lötzerich, K. *Liebigs Ann./Recueil* **1997**, 1571–1575.
8. Khanbabaee, K.; Lötzerich, K.; Borges, M.; Großer, M. *J. Prakt. Chem.* **1999**, *341*, 159–166.
9. Khanbabaee, K.; Schulz, C.; Lötzerich, K. *Tetrahedron Lett.* **1997**, *38*, 1367–1368.
10. Khanbabaee, K.; Lötzerich, K. *Eur. J. Org. Chem.* **1999**, 3079–3083.
11. (a) Schmidt, O. T.; Schmidt, D. M.; Herok, J. *Liebigs Ann. Chem.* **1954**, *587*, 67–74. (b) Schmidt, O. T.; Blinn, F.; Lademann, R. *Liebigs Ann. Chem.* **1952**, *576*, 75–84. (c) Schmidt, O. T.; Schmidt, D. M. *Liebigs Ann. Chem.* **1952**, *578*, 31–33.
12. Haddock, E. A.; Gupta, R. K.; Haslam, E. *J. Chem. Soc., Perkin Trans. 1* **1982**, 2535–2345.
13. Okuda, T.; Yoshida, T.; Hatano, T. *J. Chem. Soc., Perkin Trans. 1* **1982**, 9–14.
14. Tanaka, T.; Nonaka, G.-I.; Nishioka, I.; Miyahara, K.; Kawasaki, T. *J. Chem. Soc., Perkin Trans. 1* **1986**, 369–376.
15. Gupta, R. K.; Al-Shafi, S. M. K.; Layden, K.; Haslam, E. *J. Chem. Soc., Perkin Trans. 1* **1982**, 2525–2534.

16. Haslam, E. *Prog. Chem. Org. Nat. Prod.* **1982**, *41*, 1–46.
17. The proposed models for the transition states of the galloyl coupling reaction of 4,6-di-*O*-galloyl glucose (→4,6-*O*-HHDP ellagitannins) correlate the preferred formation of (*S*)-HHDP entities with conformational strain energy in the intermediates leading to the (*R*)-HHDP ellagitannins.[2] However, for both transition states, the *tg*-forms[22] of the glucose 6-CH_2O-galloyl residues were assumed. Although these rotamers frequently emerge from force-field calculations as the most stable ones, the *tg*-form is in fact the least populated state for D-glucose and its derivatives,[22,23] and does not play any significant role in glucose chemistry. The postulated transition states for the formation of 2,3-*O*-HHDP ellagitannins are almost equal in energy with very similar interatomic distances between the decisive galloyl carbon atoms, and strain is built up only in the final products.[2] To the best of our knowledge, no models for the formation of 3,6-*O*- or 2,4-*O*-HHDP ellagitannins have been put forth on an atomic level. The Monte-Carlo (MC) simulations described in Ref. 2 also did not allow to carry out a comprehensive geometry analysis of the complex fused ring systems of the ellagitannins.
18. Allen, F. H.; Kennard, O. *Chem. Des. Automat. News* **1993**, *8*, 1, 31–37; Cambridge Crystallographic Data File (Oct. 1999), Version 5.18. Refcodes: SAQZUF[19] and TASQOT.[20]
19. Luger, P.; Weber, M.; Kashino, S.; Amakura, Y.; Yoshida, T.; Okuda, T.; Beurskens, G.; Dauter, Z. *Acta Crystallogr., Sect. B* **1998**, *54*, 687–694.
20. Itoh, T.; Chika, J.; Shirakami, S.; Ito, H.; Yoshida, T.; Kubo, Y.; Uenishi, J. *J. Org. Chem.* **1996**, *61*, 3700–3705; erroneously the mirror image geometry of **13** is filed in the CCDF.
21. (a) Cremer, D. A.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358. (b) Jeffrey, G. A.; Taylor, R. *Carbohydr. Res.* **1980**, *81*, 182–183.
22. In carbohydrate nomenclature, the three staggered conformations of the glucopyranose 6-CH_2OH groups are denoted *gauche–gauche* (*gg*), *gauche–trans* (*gt*), and *trans–gauche* (*tg*). The first letter refers to the position of O-6 relative to O-5, the second relative to C-4; the torsion angle ω is defined through the atoms O_5–C_5–C_6–O_6. For glucose, the *gg*- and *gt*-rotamers are equally preferred over *tg*, which is destabilized by 1,3-diaxial like interactions between O-6 and O-4.[23] For the C-4 epimer galactose, the conformers are populated in the order $gt \approx tg > gg$ (Scheme 1).
23. (a) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin/New York, 1991. (b) Kroon-Batenburg, L. M. J.; Kroon, J. *Biopolymers* **1990**, *29*, 1243–1248. (c) Bock, K.; Duus, J. Ø. *J. Carbohydr. Chem.* **1994**, *13*, 513–543.
24. Yang, D.; Yip, Y.-C.; Tang, M.-W.; Wong, M.-K.; Zheng, J.-H.; Cheung, K.-K. *J. Am. Chem. Soc.* **1996**, *118*, 491–492.
25. To the best of our knowledge, D-galactose or D-mannose derived ellagitannins are not known to occur in nature; but we envisage their chemical synthesis as an interesting target.
26. The six examples of 3,6-anhydro glycosides contained in the CCDF[18] invariably display almost undistorted $^1C_4$ *anti*-chair geometries of their pyranose portions.
27. Preliminary results indicate that for the 2,2′-bismethylenbiphenyl analogs of methyl 4,6-*O*-diphenoyl-β-D-glucoside (-COO-linkages→-CH_2O-) the (*R*)-atropisomers are preferred over the (*S*)-forms.
28. Tanaka, T.; Tong, H.-H.; Xu, Y.-M.; Ishimaru, K.; Nonaka, G.-I.; Nishioka, I. *Chem. Pharm. Bull.* **1992**, *40*, 2975–2980.
29. Immel, S. *MolArch+: Molecular Architecture Modeling Program*; Darmstadt University of Technology, 2000.
30. Lindner, H. J.; Kroeker, M. *PIMM91-Closed Shell PI-SCF-LCAO-MO-Molecular Mechanics Program*; Darmstadt University of Technology, 1997. Smith, A. E.; Lindner, H. J. *J. Comput.-Aided Mol. Des.* **1991**, *5*, 235–262.

THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 276, No. 40, Issue of October 5, pp. 37415–37425, 2001
Printed in U.S.A.

Metabolism of Sucrose and Its Five Linkage-isomeric α-D-Glucosyl-D-fructoses by *Klebsiella pneumoniae*

PARTICIPATION AND PROPERTIES OF SUCROSE-6-PHOSPHATE HYDROLASE AND PHOSPHO-α-GLUCOSIDASE*

Received for publication, July 11, 2001, and in revised form, July 24, 2001
Published, JBC Papers in Press, July 25, 2001, DOI 10.1074/jbc.M106504200

John Thompson‡§, Stanley A. Robrish‡, Stefan Immel¶, Frieder W. Lichtenthaler¶, Barry G. Hall‖, and Andreas Pikis‡‡**

*From the ‡Microbial Biochemistry and Genetics Unit, Oral Infection and Immunity Branch, NIDCR, National Institutes of Health, Bethesda, Maryland 20892, the ¶Institut für Organische Chemie, Technische Universität Darmstadt, D-64287 Darmstadt, Germany, the ‖Biology Department, University of Rochester, Rochester, New York 14627-0211, the **Vaccine and Therapeutic Development Section, Oral and Infection and Immunity Branch, NIDCR, National Institutes of Health, Bethesda, Maryland 20892, and the ‡‡Department of Infectious Diseases, Children's National Medical Center, Washington, D. C. 20010-2970*

***Klebsiella pneumoniae* is presently unique among bacterial species in its ability to metabolize not only sucrose but also its five linkage-isomeric α-D-glucosyl-D-fructoses: trehalulose, turanose, maltulose, leucrose, and palatinose. Growth on the isomeric compounds induced a protein of molecular mass ~ 50 kDa that was not present in sucrose-grown cells and which we have identified as an NAD^+ and metal ion-dependent 6-phospho-α-glucosidase (AglB). The *aglB* gene has been cloned and sequenced, and AglB (M_r = 49,256) has been purified from a high expression system using the chromogenic *p*-nitrophenyl α-glucopyranoside 6-phosphate as substrate. Phospho-α-glucosidase catalyzed the hydrolysis of a wide variety of 6-phospho-α-glucosides including maltose-6′-phosphate, maltitol-6-phosphate, isomaltose-6′-phosphate, and all five 6′-phosphorylated isomers of sucrose (K_m ~ 1–5 mM) yet did not hydrolyze sucrose-6-phosphate. By contrast, purified sucrose-6-phosphate hydrolase (M_r ~ 53,000) hydrolyzed only sucrose-6-phosphate (K_m ~ 80 μM). Differences in molecular shape and lipophilicity potential between sucrose and its isomers may be important determinants for substrate discrimination by the two phosphoglucosyl hydrolases. Phospho-α-glucosidase and sucrose-6-phosphate hydrolase exhibit no significant homology, and by sequence-based alignment, the two enzymes are assigned to Families 4 and 32, respectively, of the glycosyl hydrolase superfamily. The phospho-α-glucosidase gene (*aglB*) lies adjacent to a second gene (*aglA*), which encodes an EII(CB) component of the phosphoenolpyruvate-dependent sugar: phosphotransferase system. We suggest that the products of the two genes facilitate the phosphorylative translocation and subsequent hydrolysis of the five α-D-glucosyl-D-fructoses by *K. pneumoniae*.**

The discovery in 1964 of the phosphoenolpyruvate-dependent sugar:phosphotransferase system (PEP:PTS)[1] by Roseman and colleagues (1) is a landmark in our understanding of carbohydrate dissimilation by microorganisms. Since the initial description of this multi-component system in *Escherichia coli*, the PEP:PTS has been established as the primary route for the transport and concomitant phosphorylation of a wide variety of sugars by bacteria from both Gram-negative (2, 3) and Gram-positive genera (4, 5). In many species, including *Bacillus subtilis*, *Lactococcus lactis*, *Streptococcus mutans*, *Escherichia coli*, and *Klebsiella pneumoniae* (6, 7), sucrose is accumulated via the PTS simultaneously with phosphorylation at C-6 of the glucopyranosyl moiety of the disaccharide. Intracellularly, sucrose-6-phosphate (sucrose-6-P) is hydrolyzed by sucrose-6-phosphate hydrolase (8, 9) to glucose-6-phosphate and fructose, which are then fermented via the glycolytic pathway to yield primarily lactic acid.

The structures of sucrose, its five isomeric α-D-glucosyl-D-fructoses (trivially designated trehalulose, turanose, maltulose, leucrose, and palatinose), and some related α-linked disaccharides are depicted in Fig. 1. In contrast to the many reports of sucrose fermentation, there are few references to the utilization of the isomeric glucosyl-fructoses by microorganisms (12). This fact is of particular relevance to oral biology in light of the associative role(s) of sucrose and streptococcal species in the etiology of dental caries (13, 14). Sucrose is the precursor for glucan synthesis that facilitates attachment of *S. mutans* to the tooth surface; subsequent fermentation of the disaccharide to lactic acid initiates the demineralization of tooth enamel. In this context, isomers of sucrose attract attention as potential substitutes for dietary sucrose (15–17) because they are about half as sweet as sucrose, are not metabolized (noncariogenic), and, in the case of palatinose and leucrose, are produced on an industrial scale (18, 19). From the limited information available, one might reasonably conclude that the isomers cannot be translocated by the membrane-localized transporter EII(CB) of the sucrose-PTS or that intracellular sucrose-6-P hydrolase is unable to hydrolyze the phos-

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF337811.

§ To whom correspondence and reprint requests should be addressed: NIDCR, National Institutes of Health, Bldg. 30, Rm. 528, Convent Dr. MSC-4350, Bethesda, MD 20892. Tel.: 301-496-4083; Fax: 301-402-0396; E-mail: jthompson@dir.nidcr.nih.gov.

[1] The abbreviations used are: PEP:PTS, phosphoenolpyruvate-dependent sugar:phosphotransferase system; pNPαGlc, *p*-nitrophenyl α-D-glucopyranoside; pNPαGlc6P, *p*-nitrophenyl α-D-glucopyranoside 6-phosphate; MES, 2(*N*-morpholino)ethanesulfonic acid; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; IPTG, isopropyl-β-D-thiogalactopyranoside; MD, molecular dynamics; MLP, molecular lipophilicity pattern.

phorylated PTS products.

Our interest in these issues stemmed from a survey of disaccharide utilization by *K. pneumoniae* that revealed excellent (and unexpected) growth of this organism on all five isomers of sucrose (12). Furthermore, although organisms grown previously on a particular isomer readily metabolized sucrose and all other isomers, cells of *K. pneumoniae* grown previously on sucrose fermented only sucrose (12). Comparative analyses of proteins in various cell extracts (by two-dimensional PAGE) revealed high level expression of a specific polypeptide (molecular mass ~ 50 kDa) during growth on the isomers, but this protein was not induced by growth of the organism on sucrose. These observations provided the first indication that for *K. pneumoniae*, the initial steps in metabolism of sucrose, and those of its analogs, might be separable and distinct. In the present study we have identified two adjacent genes (*aglA* and *aglB*) in *K. pneumoniae* that encode a membrane-localized transport protein of the PTS (EIICB, or AglA) and a nucleotide (NAD^+) plus metal-dependent phospho-α-glucosidase (AglB), respectively. Together, these proteins facilitate the phosphorylative translocation and subsequent hydrolysis of the five α-linked isomers of sucrose.

To facilitate the comparison of the properties of sucrose-6-P hydrolase with those of AglB, the genes encoding the two proteins (*scrB* (7) and *aglB*, respectively) have been cloned, and both enzymes have been purified from high expression systems. Recently, we prepared trehalulose-6′-P, turanose-6′-P, maltulose-6′-P, leucrose-6′-P, and palatinose-6′-P in substrate quantity (12), and the availability of these novel compounds permitted the determination of the substrate specificities of highly purified AglB and sucrose-6-P hydrolase. Remarkably, sucrose-6-P hydrolase, which by sequence-based alignment is assigned to Family 32 of glycosyl hydrolases, hydrolyzed only sucrose-6-P. In contrast AglB, which belongs to Family 4, catalyzed the cleavage of the five isomeric 6′-phosphoglucosylfructoses. In this paper, a comparative assessment of conformational, overall shape and polarity features of sucrose-6-P and its isomeric disaccharide-6′-phosphates is given, providing insight into the molecular basis for substrate discrimination by the two phosphoglucosyl hydrolases.

EXPERIMENTAL PROCEDURES

Materials—Carbohydrates were obtained from the following sources: trehalulose from Südzucker, Mannheim/Ochsenfurt, Germany; maltulose and isomaltose from TCI America; leucrose from Fluka; and palatinose from Wako Chemicals. Sucrose, turanose, and other high purity sugars were purchased from Pfanstiehl Laboratories. Maltitol, $NADP^+$, trehalose-6-P, *p*-nitrophenyl α-D-glucopyranoside (pNPαGlc), and PEP were obtained from Sigma. Phosphorylated derivatives trehalulose-6′-P, sucrose-6-P; turanose-6′-P, maltulose-6′-P, leucrose-6′-P; palatinose-6′-P, maltose-6′-P, isomaltose-6′-P, and maltitol-6-P were prepared in this laboratory by PEP:PTS activity in permeabilized (palatinose-grown) cells of *K. pneumoniae* (12). The chromogenic substrate pNPαGlc6P was prepared by selective phosphorylation of pNPαGlc with phosphorus oxychloride in trimethyl phosphate containing small proportions of water (20). Glucose-6-phosphate dehydrogenase/hexokinase (EC 1.1.1.49; EC 2.7.1.1), and phosphoglucose isomerase (EC 5.3.1.9) were from Roche Molecular Biochemicals. Ultrogel AcA-44 and TrisAcryl M-DEAE were from Sepracor.

Growth of K. pneumoniae ATCC 23357—The organism was grown at 37 °C in 1-liter bottles, each containing 800 ml of the medium defined by Sapico *et al.* (21) supplemented with 0.4% (w/v) of the appropriate sugar. After growth to stationary phase, the cells were harvested by centrifugation (13,000 × *g* for 10 min at 5 °C) and washed twice in 25 mM Tris-HCl buffer (pH 7.5) containing 1 mM $MgCl_2$. The yield was ~2 g wet weight of cells/liter.

Electrophoresis Procedures—SDS-PAGE was carried out in the Novex XCell Mini-Cell system (Invitrogen). Novex NuPage (4–12%) Bis-Tris gels and MES-SDS running buffer (pH 7.3) were used together with Novex Mark 12™ protein standards, and proteins were stained with Coomassie Brilliant Blue R-250. For Western blots, proteins were transferred to nitrocellulose membranes using NuPage transfer buffer and SeeBlue™ prestained standards. The Amersham Pharmacia Biotech Multiphor flat-bed electrophoresis unit, precast Ampholine PAG plates (pH range, 3.5–9.5) and broad range standards were used for electrofocusing experiments.

Analytical Methods—During purification, the activity of AglB in column fractions was detected by hydrolysis of the chromogenic substrate, pNPαGlc6P. The specific activity of the enzyme was determined in a discontinuous assay that contained in 2-ml: 0.1 M Tris-HCl buffer (pH 7.5), 1 mM $MnCl_2$, 0.5 mM NAD^+, and 1 mM pNPαGlc6P. After the addition of the enzyme preparation, samples of 0.25 ml were removed at 20-s intervals (over a 2-min period) and immediately injected into 0.75 ml of 0.5 M Na_2CO_3. The $A_{400\ nm}$ of the yellow solution was measured, and rates of pNP formation were calculated by assuming a molar extinction for the *p*-nitrophenoxide anion $\epsilon = 18{,}300\ M^{-1}\ cm^{-1}$. One unit of AglB activity is the amount of enzyme that catalyzes the formation of 1 μmol of pNP min^{-1}. Two-dimensional polyacrylamide gel electrophoresis (PAGE) and protein microsequencing were carried out by Kendrick Laboratories, Inc. and by the Protein Chemistry Core Facility, Columbia University, NY, respectively. The mass of AglB was determined by electrospray in an HP1100 mass spectrometer, and the sequence of N-terminal amino acids was determined with an ABI 477A protein sequencer (Applied Biosystems Inc.) with an on-line ABI 120A phenylthiohydantoin analyzer. Protein concentrations were determined by the BCA protein assay kit (Pierce). The procedure for immunodetection of AglB with polyclonal antibody to MalH from *F. mortiferum* has been described previously (20).

Cloning and Characterization of a Region Encoding the aglA and aglB Genes of K. pneumoniae ATCC 23357—Initially, using the unfinished genome sequence of *K. pneumoniae* (Washington University Genome Sequencing Center, St. Louis, MO) and our own sequence data later, five primer sets were designed to amplify, clone, and characterize the DNA fragment encoding genes *aglA* and *aglB* of *K. pneumoniae* ATCC 23357. The five primer sets were constructed as follows: KP1F-KP1R, 5′-GCCAGTTTTTTCTCTCCTGGTAGC-3′ and 5′-GCATATTACGAAAGACGGYCCAGC-3′; KP2F-KP2R, 5′-CCCTACGAGTTGTTACATGAGGATTTC-3′ and 5′-CCCCCAATGACCACAAACG-3′; KP3F-KP3R, 5′-GGCTGGACCGTCTTTCGTAATATG-3′ and 5′-TTCGAGTTACCGTGCAGGGCAAAG-3′; KP4F-KP4R, 5′-CGCTTGGGTGTGGGTTACAC-3′ and 5′-GCCGTGGTTTTACCTCGTGC-3′; KP5F-KP5R, 5′-CCCTGATCCTGCGTCTGAACC-3′ and 5′-GTTAGCCAGCGAA AAGCGG-3′.

The components of the amplification mixtures (100 μl) were: 5 units of *Pfu* DNA polymerase (Stratagene, La Jolla, CA), 1× buffer provided by the manufacturer, 20 mM each of the four DNTPs, 100 ng of DNA, and 250 ng of each primer. Amplifications were carried out in a thermal cycler (PerkinElmer 9600, PerkinElmer Life Sciences). After an initial 2-min denaturation at 95 °C, the mixtures were subjected to 30 cycles of amplification. Each cycle consisted of 1 min denaturation at 95 °C, 1 min annealing at 58 °C, and extension at 72 °C for 2 min/kilobase of insert. These were followed by a 10-min runoff at 72 °C. The PCR products were purified (QIAquick PCR purification kit, Qiagen) and ligated into pCR-Blunt vector (Invitrogen, Carlsbad, CA). After transformation into *E. coli* TOP 10 competent cells, colonies were selected on LB agar plates containing 50 μg/ml kanamycin.

Cloning of the K. pneumoniae ATCC 23357 aglB Gene in E. coli—For amplification of the gene *aglB*, two primers were synthesized from the nucleotide sequence shown in Fig. 4: forward primer KPBF, 5′-CCCACCATG**GGAGGCAGTATCATG**-3′ (the *aglB* sequence, base pairs 2001–2015, is in bold face, and the *Nco*I site is underlined); reverse primer KPBR, 5′-CCCAGAATTC**TTAATGCAGCTCAGG**-3′ (the sequence complementary to the downstream region of *aglB*, base pairs 3321–3335, is in bold face, and the *Eco*RI site is underlined). PCR amplification was performed using high fidelity *Pfu* DNA polymerase. The amplified 1.3-kilobase DNA fragment was digested with restriction endonucleases (*Nco*I and *Eco*RI), electrophoresed through 1% agarose gel, and purified (QIAquick gel extraction kit). The fragment was ligated into the similarly digested (*Nco*I-*Eco*RI) high expression vector pSE380 (Invitrogen) to form pAP-16. (In this construct, the *aglB* gene is under control of the powerful *trc* hybrid promoter, which is also regulated by the *lac*O operator and the product of the $lacI^q$ gene. Because the plasmid also carries *lacI*, expression of *aglB* is strongly repressed in the absence but is fully induced in the presence of isopropyl-β-D-thiogalactopyranoside (IPTG)). Plasmid pAP-16 was transformed into competent cells of *E. coli* TOP 10 (Invitrogen), and transformants were selected on LB agar containing 150 μg/ml ampicillin.

DNA Sequence Analysis—DNA fragments cloned in pCR-Blunt vector were sequenced by the dideoxynucleotide chain termination method

using the Sequenase 7-deaza-dGTP sequencing kit (U. S. Biochemicals, Cleveland, OH), and [α-^{35}S]deoxyadenosine triphosphate was used for labeling. For all clones, both strands of DNA inserts were sequenced. The MacVector sequence analysis package (Version 7.0, Genetics Computer Group, Madison, WI) was used to compile, edit, and analyze the results.

Growth of Cells and Preparation of Extract Containing AglB—E. coli TOP 10 (pAP-16) was grown at 37 °C in LB medium containing ampicillin (150 μg/ml) to a density $A_{600\,nm}$ ~ 0.4 units. IPTG (0.5 mM) was then added to the culture, and growth was continued for 3 h. The culture was harvested by centrifugation (13,000 × *g* for 10 min at 5 °C), and the cells (~ 2.1 g wet weight/liter) were washed by resuspension and centrifugation from 25 mM Tris-HCl buffer (pH 7.5) containing 1 mM $MnCl_2$ and 0.1 mM NAD^+ (designated TMN buffer). Washed cells (~38 g) were resuspended in 80 ml of TMN buffer, and the organisms were disrupted (at 0 °C) by 2 × 1.5-min periods of sonic oscillation in a Branson instrument (model 350) operating at ~75% of maximum power. The extract was clarified by centrifugation (180,000 × *g* for 2 h at 5 °C), and the high-speed supernatant was transferred to sacs and dialyzed overnight against 4 liters of TMN buffer.

Purification of AglB

The enzyme was purified by low pressure chromatography, and all procedures were performed in a cold room.

Step 1: TrisAcryl M-DEAE (Anion Exchange) Chromatography—Dialyzed high-speed supernatant (~85 ml) was transferred at a flow rate of 0.8 ml/min to a column of TrisAcryl M-DEAE (2.6 × 14 cm) previously equilibrated with TMN buffer. Nonadsorbed material was removed by washing with TMN buffer, and AglB was eluted with 800 ml of a linear, increasing concentration gradient of NaCl (0–0.3 M) in TMN buffer. Fractions (8 ml) were collected, and AglB activity was revealed by the intense yellow color formed upon addition of fraction samples (4 μl) to microtiter wells containing 100 μl of pNPαGlc6P assay solution. Fractions with the highest activity (22–26) were pooled and concentrated to 19 ml by pressure filtration (Amicon PM-10 membrane, 40 psi). Ammonium sulfate crystals (1.9 g) were added slowly with stirring to a concentration of 0.75 M.

Step 2: Phenyl-Sepharose CL-4B (Hydrophobic) Chromatography—The ~ 20 ml solution from step 1 was transferred (flow rate 0.5 ml/min) to a column of phenyl-Sepharose CL-4B (2.6 × 14 cm) equilibrated with TMN buffer containing 0.75 M $(NH_4)_2SO_4$. Nonadsorbed protein(s) were eluted, and then 500 ml of a decreasing, linear gradient of $(NH_4)_2SO_4$ (0.3–0 M) in TMN buffer was passed through the column. Fractions of 5 ml were collected, and AglB was recovered primarily in fractions 45–60. These fractions were pooled and concentrated by Amicon filtration to ~9 ml.

Step 3: Ultrogel AcA-44 (Molecular Sieve) Chromatography—Approximately 3 ml of preparation from step 2 was applied at a flow rate of 0.15 ml/min to a column of Ultrogel AcA-44 (1.6 × 94 cm) previously equilibrated with TMN buffer containing 0. 1 M NaCl. Fractions of 2.1 ml were collected, and those containing maximum AglB activity (50–53) were pooled. (This procedure was repeated with the remaining 2 × ~3-ml portions of concentrate from phenyl-Sepharose chromatography). AcA-44 chromatography yielded a total of ~24 ml of highly purified AglB (3 mg/ml; specific activity 4.15 units/mg).

Kinetic Analysis and Substrate Specificity of AglB—A continuous spectrophotometric assay was used for substrate specificity studies and determination of kinetic parameters for AglB. This indirect glucose-6-P dehydrogenase/$NADP^+$-coupled assay monitors formation of glucose-6-P during the AglB-catalyzed hydrolysis of substrates. The standard 1-ml assay contained: 0.1 M HEPES buffer (pH 7.5), 1 mM $MgCl_2$, 1 mM $MnCl_2$, 1 mM NAD^+, 1 mM $NADP^+$, 1 mM substrate (6′-P isomer of sucrose, or phospho-α-glucoside), and 2 units of glucose-6-P dehydrogenase/hexokinase. Reactions were initiated by addition of 15 μl (45 μg) of AglB preparation, and the increase in $A_{340\,nm}$ was recorded in a Beckman DU 640 spectrophotometer. Initial rates were determined using the kinetics program of the instrument, and a molar extinction coefficient $\epsilon = 6,220\ M^{-1}\ cm^{-1}$ was assumed for calculation of NADPH formed (equivalent to glucose-6-P liberated). In kinetic analyses the concentration range of substrate was usually 0.2–4 mM, and kinetic parameters were determined from Hofstee plots with an Enzyme Kinetics program (dogStar software, Version 1.0c). The products of turanose-6′-P hydrolysis (glucose-6-P and fructose) were determined by inclusion of 5 mM ATP and 2 units of phosphoglucose isomerase in the assay.

Cloning of the Sucrose-6-P Hydrolase Gene (scrB) from K. pneumoniae

The *scrB* gene was amplified from *K. pneumoniae* genomic DNA using the low-error-rate FailSafe™ polymerase (Epicentre). In the forward primer (5′-GGCCATGG**CGCTCTCTCTGACGCTGAA**-3′), base pairs 3119–3137 of the *K. pneumoniae scrYAB* operon (GenBank™ accession number X57401) are bolded, and the *Nco*I site is underlined. In the reverse primer (5′-GGGGGTCGAC**TACGCGTTTGGTTTTCATCA**-3′), base pairs 4587–4606 of *scrYAB* are bolded, and the *Sal*I site is underlined. The amplicon was digested with *Nco*I and *Sal*I and ligated into similarly digested pProEX Hta (Life Technologies, Inc.), and the recombinant plasmid(s) was transformed into *E. coli* K12 strain DH5α-E (Life Technologies, Inc.). Ampicillin-resistant transformants were selected, and the *scrB* genes of four plasmids containing inserts of approximately the correct size were sequenced. All shared the following differences from the published (7) *scrB* coding sequence: 991C→T, 1006G→T, 1249C→T, 1270C→T, 1549C→T, 1552A→G, 1675G→A, 1699T→C, 1738G→A (numbering as in pScrBLong). All of these differences are silent, and one plasmid was chosen and designated pScrBLong.

Growth of E. coli DH5αE (pScrBLong) and Expression of Sucrose-6-P Hydrolase

The organism was grown in LB medium containing 200 μg/ml ampicillin. At $A_{600\,nm}$ = 0.5, IPTG was added (1 mM) and growth was continued for ~ 4 h. Cells were harvested and washed with 25 mM HEPES buffer (pH 7.5) as described earlier. The yield was ~2.9 g wet weight of cells/liter.

Purification of Sucrose-6-P Hydrolase

Briefly, the purification of sucrose-6-P hydrolase was as follows. A high-speed supernatant was prepared, after resuspension and sonication, of 10 g of *E. coli* DH5αE (pScrBLong) resuspended with 20 ml of 25 mM HEPES buffer (pH 7.5). The dialyzed preparation was applied to a column of TrisAcryl M-DEAE, and after washing with the same buffer, sucrose-6-P hydrolase was eluted with an increasing gradient of NaCl (0–0.5 M). Fractions with sucrose-6-P hydrolase activity were pooled, concentrated to 8 ml, and then mixed gently with 30 ml of 0.1 M MES buffer (pH 5). Precipitated material was removed by centrifugation, and the clarified solution was applied to a column of phosphocellulose P-11 (Whatman) previously equilibrated with 0.1 M MES buffer (pH 5). Nonadsorbed proteins were removed, and sucrose-6-P hydrolase was eluted with the same buffer containing an increasing concentration of potassium phosphate buffer (0–0.1 M, pH 7). Active fractions (eluted at ~ 50 mM P_i) were pooled and concentrated. sucrose-6-P hydrolase was purified to homogeneity by passage of this solution through an AcA-44 gel filtration column previously equilibrated with 50 mM HEPES buffer, pH 7.5, containing 0.1 M NaCl. Concentration of active fractions yielded about 22 mg of sucrose-6-P hydrolase of specific activity 12.5 units/mg (with 10 mM sucrose as substrate in the assay; see below).

Sucrose-6-P Hydrolase Assay

Sucrose-6-P is the natural substrate for sucrose-6-P hydrolase, but the enzyme also hydrolyzes sucrose when the disaccharide is present at high concentration. Because of the limited availability of sucrose-6-P, the parent sugar was used as substrate during purification of sucrose-6-P hydrolase, and the glucose-6-P dehydrogenase/hexokinase-$NADP^+$-coupled assay measured glucose formed by sucrose hydrolysis. The 1-ml assay contained: 0.1 M HEPES buffer (pH 7.5), 1 mM $MgCl_2$, 1 mM $NADP^+$, 1 mM ATP, 10 mM sucrose, 2 units of glucose-6-P dehydrogenase/hexokinase, and enzyme solution. One unit of sucrose-6-P hydrolase activity is the amount of enzyme that catalyzes the formation of 1 μmol of glucose/min.

Computational Methods

A conformational analysis of all disaccharides and their 6′-phosphates was carried out using a molecular dynamics (MD) simulations, *i.e.* CHARMM (22, 23), with a force field particularly adapted for the treatment of carbohydrates (24, 25), with the explicit incorporation of water as the solvent. The starting structures used were derived either from the corresponding x-ray-based solid-state geometries of sucrose (26, 27), β-*p*-turanose (28), β-*p*-leucrose (29), β-*f*-palatinose (30), α, α-trehalose (31), β-*p*-maltose (32, 33), and maltitol (34, 35) or from compounds of similar backbone structure found in the Cambridge Crystallographic Database (www.ccdc.com.ac.uk) (36, 37). Any water molecules present in the crystal structures were removed. For compounds

that equilibrate between different anomeric or ring (pyranoid or furanoid) forms, only the most predominant tautomer was considered, *i.e.* the 6′-phosphates of β-*p*-trehalulose, β-*p*-turanose, β-*p*-maltulose, β-*p*-leucrose, β-*f*-palatinose, β-*p*-maltose, and β-*p*-isomaltose. Each compound was centered in a periodic box (truncated octahedron, box size ~ 33.5 Å) filled with pre-equilibrated TIP3 (transferable intermolecular potential-3) water molecules, yielding (after removal of the solvent molecules that overlap with the solute) simulation systems including 643 (disaccharides) or 641 (disaccharide phosphates) water molecules, respectively. In the latter series, two NH_4^+ counterions were added at random positions within 6 Å around the glucose-6-$CH_2OPO_3^{2-}$ groups. After full lattice energy minimization, all boxes were slowly heated from 0 to 300 K within 15 ps of MD simulation and were subsequently equilibrated for an additional 85 ps; the final MD data were sampled using simulations of 1 ns in each case; molecular configurations were saved every 100 fs for analysis purposes. All MD runs were carried for constant pressure (P_{ref} = 1 atm, isothermal compressibility $4.63 \cdot 10^{-5}$ atm^{-1}, pressure coupling constant τ_P = 5 ps) and constant temperature (T_{ref} = 300 K, temperature coupling constant τ_T = 5 ps, allowed temperature deviation $\Delta T = \pm$ 10 K) conditions (NPT ensemble (constant number of molecules, constant pressure, and constant temperature)) using the following simulation parameters: time step Δt = 1 fs (leapfrog integrator, all *X*–H bond lengths were constrained using the SHAKE protocol), dielectric constant ϵ = 1.0, cut-off distance for long range interactions 12 Å, cut-off for images in atom lists 13 Å. The following averages were recalculated from the final MD runs (standard deviations are in parentheses): disaccharides: temperature $<T>$ = 296(5) K, box size 33.66(8) Å, volume $<V>$ = 19060(145) $Å^3$, density $<\rho>$ = 1.039(8) g cm^{-3}; disaccharide 6′-phosphates: temperature $<T>$ = 296(5) K, box size 33.59(5) Å, volume $<V>$ = 18950(130) $Å^3$, density $<\rho>$ = 1.049(8) g cm^{-3}. For each MD time series a mean solute geometry was obtained by three-dimensional fitting of all configurations (heavy atoms only, excluding CH_2OH oxygen atoms); the best-fit models from this procedure were selected as representative molecular geometries in aqueous solution (Fig. 9, 10). For a comparison, the conformation of all glucosyl-6′-CH_2OH groups were set to *gauche-trans* (*gt*, torsion angle O5′–C5′–C6′–O6′ ω = +60°). Solvent-accessible surfaces (38) and color-coded molecular lipophilicity patterns (MLPs) (39, 40) were generated using the MOLCAD modeling program (41, 42).[2]

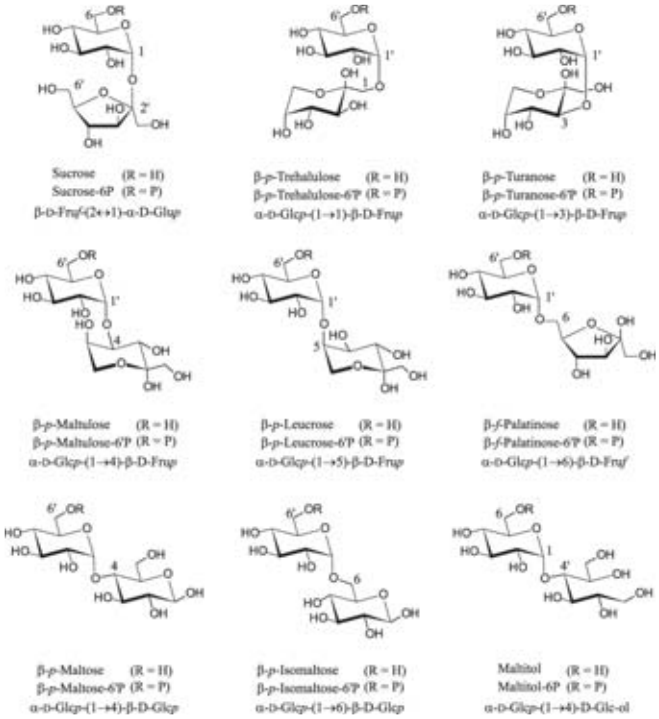


FIG. 1. **Chemical formulae and established abbreviations (10) of sucrose, its five linkage isomeric glucosyl-fructoses, and of some related disaccharides (*R* = H) and their respective monophosphates (*R* = PO_3^{2-}), invariably carrying their phosphate ester groups attached to the glucosyl-C-6.** For the reducing disaccharides, only the tautomeric form predominating in solution (10, 11) is depicted. The nonreducing sucrose-6-P is the singular substrate for the sucrose-6-P hydrolase, whereas all others are hydrolyzed by the 6-phospho-α-glucosidase described herein.

RESULTS

Growth of K. pneumoniae on Sucrose and Its Isomers—Recently (12) we reported the growth of *K. pneumoniae* on sucrose and its five isomeric α-D-glucosyl-D-fructoses (see Fig. 1 for structures). Additionally, we showed that organisms grown previously on a particular isomer readily fermented sucrose as well as each of the α-D-glucosyl-D-fructoses, whereas sucrose-grown cells, surprisingly, metabolized only sucrose (12). Examination of the protein composition of the various cell extracts by two-dimensional PAGE revealed high level expression of one specific polypeptide (molecular mass ~50 kDa) during growth on either of the five sucrose isomers (*e.g.* palatinose and maltulose, Fig. 2) and in fact on related disaccharides such as maltose, isomaltose, maltitol (for formulae, see Fig. 1), and even methyl-α-D-glucopyranoside (data not shown). Significantly, the ~50-kDa protein was not detectable in an extract prepared from sucrose-grown cells (Fig. 2).

Identity of the Protein Induced during Growth on Sucrose-Isomeric Glucosyl-fructoses—Proteins from a duplicate two-dimensional PAGE gel of maltulose-grown cell extract were transferred by Western blot to a polyvinylidene difluoride membrane. Microsequence analysis provided the following sequence for the first 25 residues from the N terminus of the highly expressed ~50-kDa protein: MKKFSVVIAGGGSTFTPGIVLMLLA. A BLAST (43) search of the nonredundant protein data bases with this sequence as probe revealed 91 and 82% identity, respectively, with the N termini of an unusual 6-phospho-α-glucosidase (EC 3.2.1.122), previously purified from *Fusobacterium mortiferum* (MalH (44)) and *B. subtilis* (GlvA (45)). Phospho-α-glucosidase activity is readily detected by the intensely yellow *p*-nitrophenolate (pNP) anion released upon hydrolysis of pNPαGlc6P. This chromogenic substrate was rapidly hydrolyzed by extracts of cells grown on the glucosyl-fructoses and other α-glucosides, but essentially no activity was detectable in the extract from sucrose-grown cells (Table I). Western blots performed with antibody raised against phospho-α-glucosidase from *F. mortiferum* (20) revealed a striking correlation between the amount of induced immunoreactive protein of ~50 kDa (Fig. 3) and the hydrolytic activities of the various extracts (Table I). The protein induced during growth on the five α-D-glucosyl-D-fructoses was thus identified as phospho-α-glucosidase.

Cloning and Sequence Analysis of the Agl Region of K. pneumoniae—Although suggestive, the available data (Figs. 2 and 3 and Table I) did not establish a functional role for phospho-α-glucosidase in dissimilation of the five α-D-glucosyl-D-fructoses by *K. pneumoniae*. Recently, we demonstrated the PEP-dependent phosphorylation of the five sucrose isomers via the PTS activity of palatinose-grown cells of *K. pneumoniae*, and trehalulose-6′-P, turanose-6′-P, maltulose-6′-P, leucrose-6′-P, and palatinose-6′-P were prepared in 20–50-mg amounts (12). To determine whether these derivatives were hydrolyzed by AglB, it was first necessary to purify this enzyme. To this end, *aglB*, the gene encoding the phospho-α-glucosidase, and an adjacent upstream gene, *aglA*, were cloned and sequenced. Fig. 4 shows the ~3.5-kilobase pair nucleotide sequence containing the two genes (*aglA* and *aglB*) of the alpha-glucoside utilization region of the *K. pneumoniae* genome. The *aglA* gene comprises a coding sequence of 1,619 nucleotides commencing with an ATG codon at position 394 and terminating with a TGA (stop) codon at position 2014. This open reading frame encodes a

[2] The major part of the MOLCAD program is included in the SYBYL package of TRIPOS Associates, St. Louis, MO.

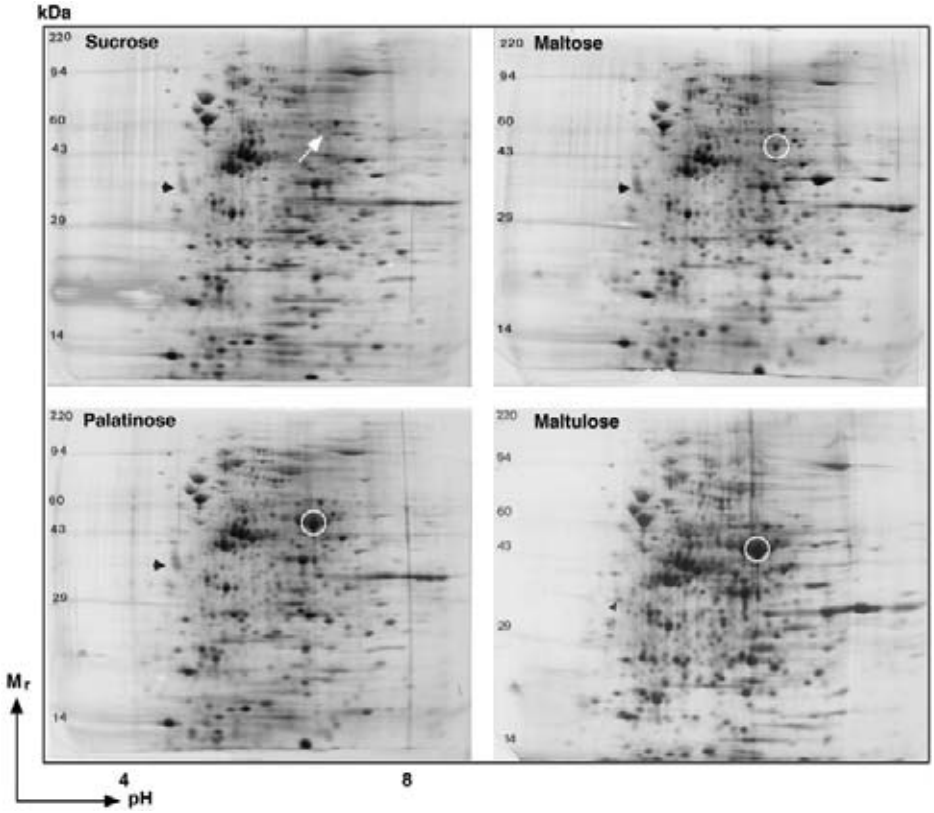


FIG. 2. **Analysis by two-dimensional PAGE of proteins in extracts prepared from cells of *K. pneumoniae* grown on various disaccharides.** The *white circles* indicate the induced ~50-kDa phospho-α-glucosidase (AglB) in organisms grown previously on either maltose, palatinose, or maltulose. This protein was not detectable (*white arrow*) in the extract prepared from sucrose-grown cells of *K. pneumoniae*. Approximately 50 μg of protein was applied per gel, and polypeptides were visualized by silver staining. Prior to electrophoresis, tropomyosin (1 μg) was added to each sample as an IEF internal standard. This protein (*black arrowhead*) migrates as a doublet with the lower polypeptide spot ~33 kDa and pI = 5.2.

TABLE I
Hydrolysis of p-nitrophenyl α-D-glucopyranoside 6-phosphate by extracts prepared from cells of K. pneumoniae grown on different sugars

Growth sugar	Rate of pNPαGlc6P hydrolysis[a]
Maltulose (α,1–4)[b]	0.256 (100)[c]
Palatinose (α,1–6)	0.214 (84)
Leucrose (α,1–5)	0.202 (79)
Trehalulose (α,1–1)	0.160 (63)
Methyl-α-D-glucoside	0.119 (46)
Turanose (α,1–3)	0.096 (38)
Maltitol	0.073 (29)
Maltose	0.031 (12)
Sucrose	0.002 (<1)
Cellobiose	0.002 (<1)
Galactose	ND[d]
Trehalose	ND
Glucose	ND

[a] μmol pNPαGlc6P hydrolyzed min^{-1} mg $protein^{-1}$.
[b] Compounds in boldface are sucrose isomers.
[c] Values in parentheses = rate as % activity in maltulose extract.
[d] ND, no detectable activity.

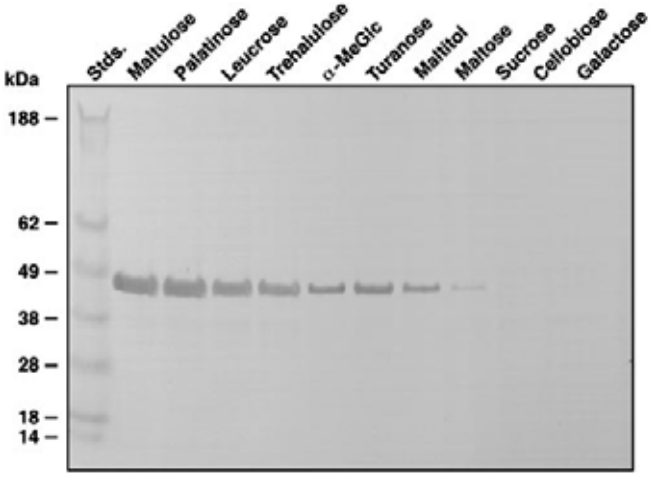


FIG. 3. **Western blot showing the sugar-specific induction and cross-reactivity of the ~50-kDa protein (AglB) with antibody raised against purified MalH (phospho-α-glucosidase) from *F. mortiferum* (20).** Extracts were prepared from cells of *K. pneumoniae* grown on the indicated sugars, and approximately 15 μg of protein was applied per lane. Note the absence of immunoreactive protein in sucrose-grown cells. *Stds.*, standards.

polypeptide of 540 residues (calculated M_r = 58,373) that contains fused C and B domains characteristic of a membrane-localized EII(CB) transport protein of the PTS (46). The *aglA* gene is preceded by a potential ribosome-binding site (GAGGA) centered ~11 nucleotides from the start codon. The *aglA* stop codon overlaps the Met start codon of *aglB*. The latter gene extends from nucleotide 2013 and terminates with a TAA codon at position 3333. Translation of *aglB* predicts a polypeptide of 440 amino acids (calculated M_r = 49,256), in which residues 138–169 display the signature pattern of Family 4 glycosyl hydrolases (47, 48):[3] P*X*(S/A)(W/T)(L/I/V/M/F)^2(Q/N)*X*^{2}NP*X*^{4}(T/A)*X* ^{9,10}(K/R)*X*(L/I/V)(G/N)*X*C. From the alignment shown in Fig. 5, it is clear that AglB exhibits homology with phospho-α-glucosidases from other species including MalH from *F. mortiferum* (75% identity), GlvA from *B. subtilis* (72%), and truncated GlvG from *E. coli* (77%), respectively.

Purification of Phospho-α-glucosidase (AglB) from E. coli TOP(pAP-16)—Cells of *E. coli* TOP(pAP-16) produced high levels of an IPTG-inducible protein with an estimated M_r ~ 50 kDa as expected for the full-length polypeptide encoded by *aglB* (Fig. 6*A*, *lane 1*). This protein cross-reacted with phospho-α-glucosidase antibody (Fig. 6*B*, *lane 1*), and the cell extract catalyzed the immediate hydrolysis of pNPαGlc6P. AglB was purified by conventional low-pressure chromatography, and to stabilize the enzyme, 0.1 mM NAD^+ and 1 mM Mn^{2+} ion were included in all buffers. Throughout the four-stage procedure,

[3] On line at www.expasy.ch/cgi-bin/lists?glycosid.txt.

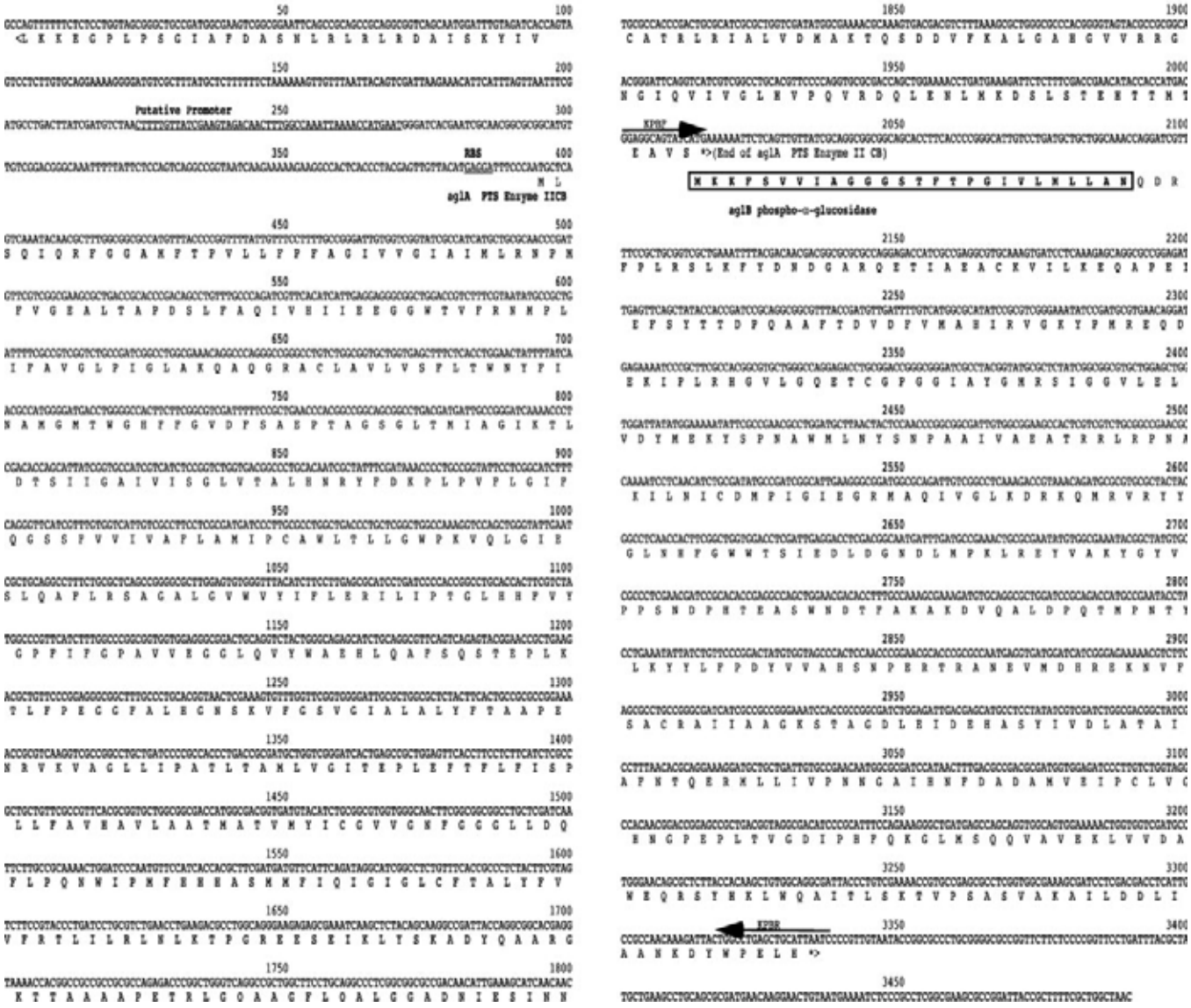


FIG. 4. **Nucleotide sequence of the Agl region of *K. pneumoniae*.** This ~3.5-kilobase DNA fragment contains genes *aglA* and *aglB* that encode an EIICB transport protein of the PEP:PTS and an NAD^+ plus metal-dependent 6-phospho-α-glucosidase, respectively. A potential ribosomal binding site (*RBS*) preceding *aglA* is *underlined*. The deduced amino acid sequences are shown *below* the nucleotide sequence in *single-letter code*. The N-terminal amino acid sequence of AglB obtained by Edman degradation is *boxed*. The positions of primers KPBF and KPBR used for PCR amplification of *aglB* are indicated by *arrows above* the nucleotide sequence.

the purification of AglB was monitored by enzymatic assay (Table II), SDS-PAGE (Fig. 6*A*), and immunoblot methods (Fig. 6*B*). Approximately 70 mg of electrophoretically pure enzyme was obtained from ~38 g wet weight of cells. Although purified in reasonably active form, AglB was progressively inactivated throughout the purification, and the specific activity of the final preparation (4.2 units/mg) was only ~3-fold higher than that of the original dialyzed cell extract (1.2 units/mg).

Properties of AglB—The molecular weight of AglB determined by electrospray/MS (M_r 49,254) was within two units of the theoretical weight average M_r of 49,256 deduced from the amino acid sequence encoded by *aglB*. However, in the final stage of purification, AglB emerged from the AcA-44 gel filtration column in a volume suggestive of a protein of molecular mass ~ 100 kDa. Cross-linking studies also revealed the formation of a similarly sized product after incubation of the enzyme with various homo-bifunctional imidoesters (Fig. 6*C*, *lanes 2–4*). It appears likely that in solution AglB exists as a catalytically active homodimer. Analytical electrofocusing revealed two species (Fig. 6*D*, *lane 2*) having estimated pI values of 5.4 and 5.6 that agreed fairly well with the theoretical pI (5.69) deduced from the amino acid composition of AglB. The homogeneity of the purified enzyme was confirmed by the unambiguous determination of the first 26 residues from the N terminus, MKKFSVVIAGGGSTFTPGIVLMLLAN. This sequence was precisely that deduced by translation of *aglB* and, importantly, was in perfect agreement with that of the polypeptide induced during growth of *K. pneumoniae* on the sucrose-isomeric glucosyl-fructoses (Fig. 2).

Cofactor, Metal Ion Requirements, and Substrate Specificity of AglB—Phospho-α-glucosidases MalH and GlvA from *F. mortiferum* (20, 44) and *B. subtilis* (45), respectively, exhibit requirements for nucleotide (NAD^+) and divalent metal ion (Mn^{2+}, Co^{2+}, or Ni^{2+}) for activity. AglB exhibited similar requirements and, in the absence of these cofactors, was unable to hydrolyze pNPαGlc6P (Table III). Inclusion of NAD^+ in the assay elicited substrate cleavage, but enzyme activity increased 3–6-fold upon further addition of Mn^{2+}, Co^{2+}, or Ni^{2+}. Other divalent metal ions tested, including Mg^{2+}, Ca^{2+}, and Zn^{2+}, were either without effect or were inhibitory. The activity

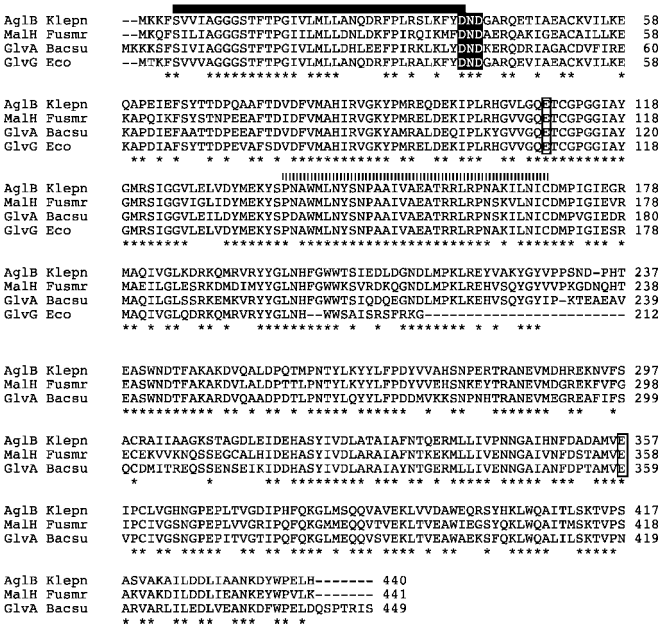


FIG. 5. **Comparative alignment of the sequence of AglB from *K. pneumoniae* with 6-phospho-α-glucosidase(s) from *F. mortiferum* 25557 (*MalH* (*Fusmr*), Swiss Protein Database accession no. O06901); *B. subtilis* 168 (*GlvA* (*Bacsu*), SwissProt identifier P54716) and *E. coli* K-12 MG1655 (*GlvG*, truncated (*Eco*) Ref. 49, Swiss Protein Database accession no. P31450).** The *bold overline* at the N terminus indicates a probable NAD^+-binding domain, and the DND motif is *highlighted*. Catalytically important glutamyl residues are *boxed*, and conserved residues are indicated by *asterisks*. Residues representing the signature motif for Family 4 glycosyl hydrolases (Swiss Protein Databank; Prosite name, PS01324 (48)) are indicated by the *shaded overline*.

of AglB was optimal at ~36 C° in either 0.1 M Tris-HCl or HEPES buffers (pH 7.5) containing 0.1 mM NAD^+ and 1 mM Mn^{2+} ion. In the presence of requisite cofactors, AglB hydrolyzed all 6-phospho-α-D-glucosides tested including all phosphorylated isomers of sucrose. The kinetic parameters for each substrate are presented in Table IV. There was no detectable cleavage of the corresponding nonphosphorylated compounds. Importantly, sucrose-6-P itself was not hydrolyzed by AglB nor was it an inhibitor of enzyme activity. Studies with turanose-6′-P (Table V) established that, as for the chromogenic analog (pNPαGlc6P), the same cofactors were required for the hydrolysis of this PTS product. Throughout the time course of the experiment, the 1:1 stoichiometry between [glucose-6-P:fructose] confirmed these two metabolites as the only reaction products from AglB-catalyzed hydrolysis of turanose-6′-P.

Purification and Substrate Specificity of Sucrose-6-P Hydrolase—AglB readily hydrolyzed the five 6-phosphoglucosyl-fructoses, whereas sucrose-6-P, remarkably, was not a substrate for this enzyme. It was of interest, to determine whether sucrose-6-P hydrolase would exhibit the converse specificity with respect to potential substrates. sucrose-6-P hydrolase was purified from *E. coli* DH5αE (pScrB Long) as described under "Experimental Procedures." The four-stage procedure (Fig. 7) provided 20–30 mg of electrophoretically pure sucrose-6-P hydrolase with an estimated molecular mass of ~53 kDa by SDS-PAGE (Fig. 7, *lane 4*), which was in agreement with the molecular weight of 52,708 deduced by translation of the *scrB* gene (ref. 7 and Swiss Protein Database accession no. P27217). The mass of sucrose-6-P hydrolase determined experimentally by electrospray mass spectroscopy (M_r 52,581) was about 127 mass units lower than the calculated $mass_{av}$. Except for the absence of methionine at the N terminus, microsequence analysis confirmed exactly the predicted sequence of the first 28 residues of the polypeptide SLPSRLPAILQAVMQGQPQALADSHYPQ. Sucrose-6-P hydrolase catalyzed the hydrolysis of sucrose and sucrose-6-P at comparable rates ($V_{max\cdot sucrose}$ = 31.2 ± 1.1; $V_{max\cdot S6P}$ = 40.4 ± 2.3 μmol hydrolyzed min^{-1} mg^{-1}). However, the affinity of the enzyme for the phosphorylated disaccharide ($K_{m\ S6P}$ = 85.3 ± 15.1 μM) was >200-fold greater than for sucrose ($K_{m\ sucrose}$ = 20.3 ± 1.9 mM). There was no detectable hydrolysis (at 1 mM) of any of the phosphorylated isomers of sucrose, and sucrose-6-P hydrolase failed to hydrolyze other phospho-α-glucosides including maltose-6′-P and trehalose-6-P.

Conformational Analysis of Sucrose-6-P and Disaccharide-6′-phosphates—Insight into the remarkable discrimination of sucrose-6-P hydrolase and phospho-α-glucosidase for their substrates was provided by conformational analysis of these phosphorylated compounds using molecular dynamics simulations. Sucrose and sucrose-6-P differ from other disaccharides not only by the fact that they are nonreducing (the two sugar units are linked through their anomeric centers) but by the predetermined orientation of the glucose and fructose portions toward each other. In the solid state, the two sugars are conformationally fixed by two intramolecular hydrogen bonds (Fig. 8 and Refs. 26, 27, 50, 51). On dissolution of the disaccharide in water, these bonds are replaced by an H_2O molecule bridging glucosyl-O-2 and fructosyl-O-1 through hydrogen bonding (Fig. 8, *center*, and Ref. 52), to yield an overall conformation close to that in the crystalline state. The molecular geometry of sucrose-6-P in water, which emerges from a nanosecond molecular dynamics simulation in a truncated octahedron box containing 641 water molecules, again is very similar to that of sucrose in the crystal and in aqueous solution (Fig. 8, *right*), so that a water bridge of the Glc-2-O···H_2O···O-1-Fru is likewise to be inferred. A comparison of the molecular geometry of sucrose-6-P in water with the geometries of the nine disaccharide-6′-phosphates reveals their distinctly different molecular shapes. Unlike sucrose-6-P, which by virtue of the intramolecular water bridge between glucose and fructose assumes a remarkably compact conformation in solution, the nine disaccharide-6′-phosphates lack any interaction of this type and hence invariably adopt a more extended, longish molecular geometry. These differences in molecular shape are emphasized by juxtaposition of the solvent-accessible surface of sucrose-6-P (Fig. 9, *top*) with those of the nine disaccharide-6′-phosphates shown superimposed in Fig. 9 (*bottom*).

DISCUSSION

Transport and Hydrolysis of Sucrose and Its Isomers by K. pneumoniae—Circumstantial evidence indicated that the transport and dissimilation of the five *O*-α-linked isomers of sucrose by *K. pneumoniae* occurred by a route different from the PTS-sucrose-6-P hydrolase route used for sucrose itself (12). For example, sucrose-grown cells failed to metabolize any of the isomers, and the PEP:PTS activity of cells grown on a particular isomer (*e.g.* palatinose) catalyzed the phosphorylation of all other isomers. Importantly, growth of *K. pneumoniae* on the five α-D-glucosyl-D-fructoses induced a high level expression of a polypeptide (molecular mass ~ 50 kDa) that was not present in organisms grown on sucrose. In this study, the gene (*aglB*) that encodes the induced protein has been cloned and sequenced, and the protein itself (AglB) has been identified as an NAD^+ and metal-dependent phospho-α-glucosidase. The gene *aglB* lies adjacent to a second gene, *aglA*, which encodes an EII(CB) component of the PEP:PTS (46). It is our contention that together, AglA and AglB facilitate the phosphorylative translocation and subsequent cleavage of phosphorylated isomers of sucrose (and related α-glucosides) by *K. pneumoniae*.

Properties of sucrose-6-P hydrolase and Phospho-α-glucosi-

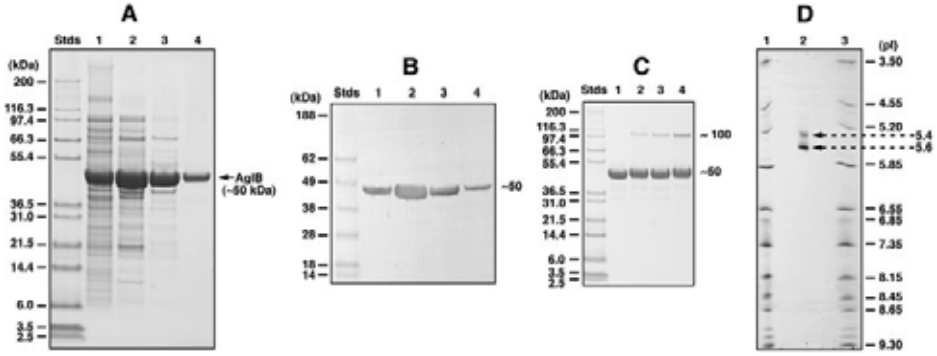


FIG. 6. **Determination of the M_r, pI, and structural composition of AglB by analytical PAGE.** *A*, purification and M_r estimate of AglB. Samples from each stage of purification were denatured, resolved by SDS-PAGE, and stained with Coomassie Brilliant Blue R-250. *Lane 1*, high-speed supernatant; *lane 2*, TrisAcryl M-DEAE; *lane 3*, phenyl-Sepharose Cl-4B; and *lane 4*, Ultrogel AcA-44. *B*, Western blot of a duplicate gel of *panel A* showing cross-reaction of AglB with MalH antibody**.** *Stds*, standards. *C*, cross-linking of AglB subunits to the dimeric state by treatment with: *lane 1*, no agent; *lane 2*, dimethyladipimidate; *lane 3*, dimethylpimelimidate; and *lane 4,* dimethylsuberimidate. *D*, determination of the pI of AglB by analytical electrofocusing (*lane 2*).

TABLE II
Summary of the purification of AglB (phospho-α-glucosidase) from E. coli TOP 10 (pAP-16)

Purification step	Total protein	Total activity	Specific activity	Purification	Yield
	mg	*units*[a]	*units/mg*	*-fold*	*%*
Dialyzed high-speed supernatant	2304	2815	1.22	1	100
TrisAcryl M-DEAE	618	1743	2.82	2.3	62
Phenyl-Sepharose CL-4B	174	715	4.12	3.4	25
AcA-44	72	299	4.15	3.4	11

[a] Units expressed as μmol of pNPαGlc6P hydrolyzed min^{-1}.

TABLE III
NAD^+ and metal ion requirements for activity of AglB (phospho-α-glucosidase)

The purified enzyme had been dialyzed against 25 mM Tris-HCl buffer (pH 7.5).

Addition to assay[a]	Specific activity[b]
No additions	ND[c]
+ NAD^+	0.35
+ Mg^{2+}	0.01
+ NAD^+ + Mg^{2+}	0.32
+ Ni^{2+}	0.11
+ NAD^+ + Ni^{2+}	0.59
+ Co^{2+}	0.26
+ NAD^+ + Co^{2+}	1.01
+ Mn^{2+}	0.25
+ NAD^+ + Mn^{2+}	1.35

[a] The 2-ml assay solution contained 50 mM Tris-HCl buffer (pH 7.5). When required, appropriate divalent metal ion (1 mM) and NAD^+ (0.5 mM) were included. Phospho-α-glucosidase (60 μg) was added, and after 2 min of preincubation at 25 °C, pNPαGlc6P was added to a final concentration of 1 mM. Hydrolysis of substrate (*i.e.* formation of pNP) was followed as described under "Experimental Procedures."

[b] Expressed as μmol of pNPαGlc6P hydrolyzed min^{-1} mg $enzyme^{-1}$.

[c] ND, no detectable activity.

TABLE IV
Substrate specificity and kinetic parameters of purified AglB (phospho-α-glucosidase) from Klebsiella pneumoniae

Compounds in bold face are sucrose isomers.

Phospho-α-glucoside	K_m	V_{max}
	mM	*μmol mg^{-1} min^{-1}*
Trehalulose-6′-P (α,1–1)	1.23 ± 0.10	0.89 ± 0.03
Turanose-6′-P (α,1–3)	1.68 ± 0.26	2.41 ± 0.19
Maltulose-6′-P (α,1–4)	1.20 ± 0.12	1.15 ± 0.05
Leucrose-6′-P (α,1–5)	5.63 ± 1.24	0.85 ± 0.13
Palatinose-6′-P (α,1–6)	2.42 ± 0.39	0.90 ± 0.08
Maltose-6′-P	3.08 ± 0.49	1.31 ± 0.13
Isomaltose-6′-P	4.48 ± 1.00	1.55 ± 0.23
Maltitol-6-P	0.82 ± 0.23	1.87 ± 0.21
Trehalose-6-P	1.16 ± 0.30	0.31 ± 0.03
pNPαGlc6P	0.05 ± 0.01	2.42 ± 0.25

TABLE V
Product stoichiometry and requirement(s) for NAD^+ and Mn^{2+} ion for hydrolysis of turanose-6′-P by AglB (phospho-α-glucosidase)

The 1-ml reaction mixture contained 1 μmol of turanose-6′-P. For composition of the reaction, details of sampling, and enzymatic analyses, see "Experimental Procedures."

Reaction time	NAD^+ and Mn^{2+} present[a]		NAD^+ and Mn^{2+} omitted	
	Glucose-6-P	Fructose	Glucose-6-P	Fructose
min				
0	ND[b]	ND	ND	ND
2	0.32[c]	0.32	0.03	0.03
5	0.57	0.53	0.05	0.05
10	0.87	0.85	0.07	0.07
15	0.95	0.91	0.10	0.08
20	1.00	0.94	0.11	0.10

[a] NAD^+ and Mn^{2+} were both present at a concentration of 1 mM.

[b] ND, not detectable.

[c] Numerical values indicate μmol of glucose-6-P and fructose formed.

dase (AglB)—In some of their properties, sucrose-6-P hydrolase and AglB show similarity. For example they are of comparable (monomer) size, they are exacting for the glucose-6-P moiety of their substrates, and both exhibit poor or no affinity for non-phosphorylated disaccharides. However, in their amino acid sequences, cofactor requirements, and assignments to different families of the glycosyl hydrolase superfamily, sucrose-6-P hydrolase and AglB are quite different. The amino acid sequence of sucrose-6-P hydrolase (deduced from the *scrB* gene (7)) has essentially no homology with that of AglB, and by the amino acid-based sequence classification of Henrissat (47), AglB and sucrose-6-P hydrolase are assigned to Families 4 and 32, respectively, of the glycosyl hydrolase superfamily (48).[3] Sucrose-6-P hydrolase has no cofactor requirements, whereas AglB is dependent upon both NAD^+ and divalent metal ion (Mn^{2+}, Ni^{2+}, or Co^{2+}) for catalytic activity (Table III). Indeed, these cofactor requirements for AglB were predicted by virtue of the extraordinarily high sequence identity between the putative polypeptide encoded by *aglB* and those of the Family 4 phospho-α-glucosidases shown in the multiple alignment in Fig. 5. The role(s) for NAD^+ and metal ion have not been established,

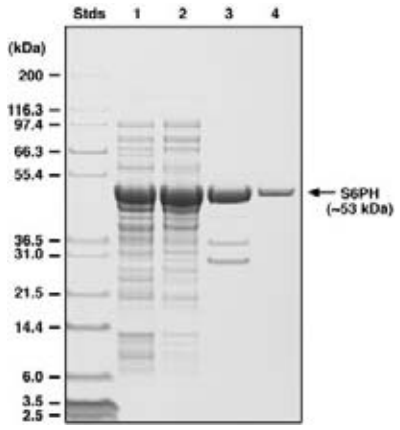


FIG. 7. **SDS-PAGE of samples from each stage of purification of sucrose-6-P hydrolase from *E. coli* DH5αE (pScrBLong).** *Lane 1*, high-speed supernatant; *lane 2*, TrisAcryl M-DEAE; *lane 3*, phosphocellulose P-11; and *lane 4*, Ultrogel AcA-44 gel filtration chromatography. *Stds*, standards.

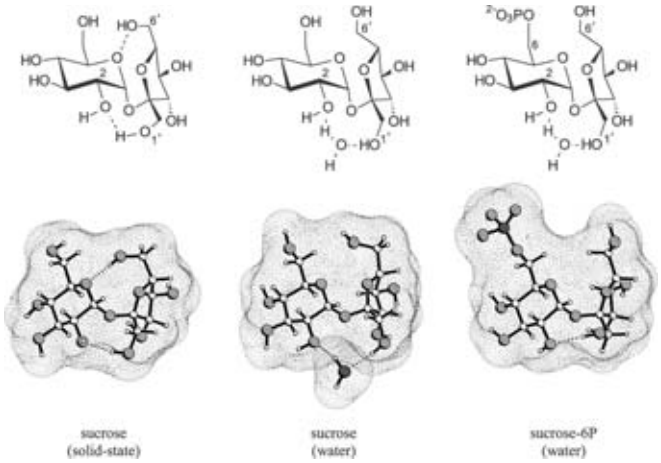


FIG. 8. **Preferred molecular geometries of sucrose in the solid state (*left*) characterized by two intramolecular hydrogen bonds (26, 27, 50, 51) and in water (*center*), shown here with the water molecule bridging glucosyl-O-2 and fructosyl-O-1 through hydrogen bonding (52).** The conformation of sucrose-6-phosphate emerging from a 1000-ps MD simulation in a box containing 641 water molecules (*right*) is so similar to that of sucrose in water that a Glc-2-O···H_2O···O-1-Fru water bridge is likewise inferred. The *dotted contours* refer to the solvent-accessible surface into which *ball-and-stick* models have been inserted; for easier comparison, the glucosyl moiety is kept in the same orientation.

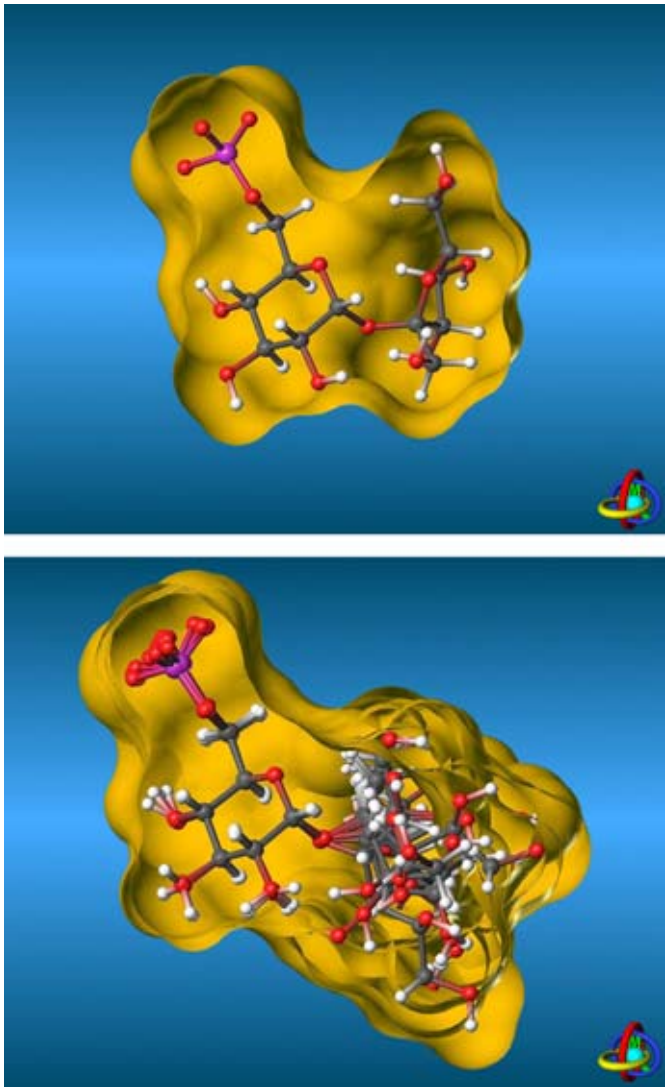


FIG. 9. **Solvent-accessible surface of sucrose-6-P in front-opened form with *ball-and-stick* model insert (*top*) as set against those of the nine other disaccharide-6′-phosphates (*bottom*) superimposed on each other with the α-D-glucose-6-P portion (*left half*) kept in the same orientation.** The slender form of the fructose moiety of sucrose-6-P (*top, right half*) renders the shape of the molecule different; notably it is more compact than that of the other disaccharide-6-phosphates.

and it is presently unclear whether these cofactors play catalytic and/or structural roles in AglB and related enzymes of Family 4 (44, 45, 53, 54). Results obtained from site-directed mutagenesis of the phospho-α-glucosidase (GlvA) in *B. subtilis* (45) suggest that residues close to the N terminus comprise the NAD^+-binding domain (*see*, Fig. 5). Glycosyltransferases comprise a superfamily of Mn^{2+}-dependent enzymes (55) that use UDP-glucose, UDP-galactose, and related compounds as substrates for modification (via glycosylation) of a wide variety of biological molecules in both prokaryotes and eukaryotes. Most, if not all, members of this large family contain a conserved motif D(*X*)D that participates in the substrate recognition/catalytic process by interaction of the aspartyl residues with the ribose moiety of the nucleotide or via coordination with Mn^{2+} ion (56). Interestingly, this motif is also present in AglB and in other phospho-α-glucosidases, and the conserved DND residues lie adjacent to the putative NAD^+-binding domain of these enzymes (Fig. 5). Furthermore, site-directed substitution at the first aspartic residue of this motif (D41G and D41E) in GlvA results in loss of hydrolytic activity (45). These findings plus the fact that the D(*X*)D motif is conserved in other members of Family 4 (see Fig. 4 in Ref. 45) may indicate a role for the two acidic residues in Me^{2+} ion-binding in AglB and related glycosyl hydrolases.

Substrate Discrimination by Sucrose-6-P Hydrolase and 6-Phospho-α-glucosidase—Sucrose-6-P and its five phosphorylated linkage isomers have recently been prepared and characterized by ^1H and ^{13}C NMR spectroscopy (12). The availability of these derivatives in substrate amount permitted specificity and kinetic analyses to be carried out with highly purified sucrose-6-P hydrolase and AglB. These studies establish unequivocally that sucrose-6-P hydrolase hydrolyzes only sucrose-6-P to form glucose-6-P and fructose. The specificity of sucrose-6-P hydrolase for its single substrate (sucrose 6-phosphate) is noteworthy because it suggests that their reciprocal molecular recognition (as a prerequisite to fission of the intersaccharidic linkage to glucose-6-P and fructose) is unique, not even tolerating minor changes in the linkup of the two sugars, as for example those realized in the five isomeric glucosyl-fructoses. In contrast, the 6-phospho-α-glucosidase (AglB), which is in-

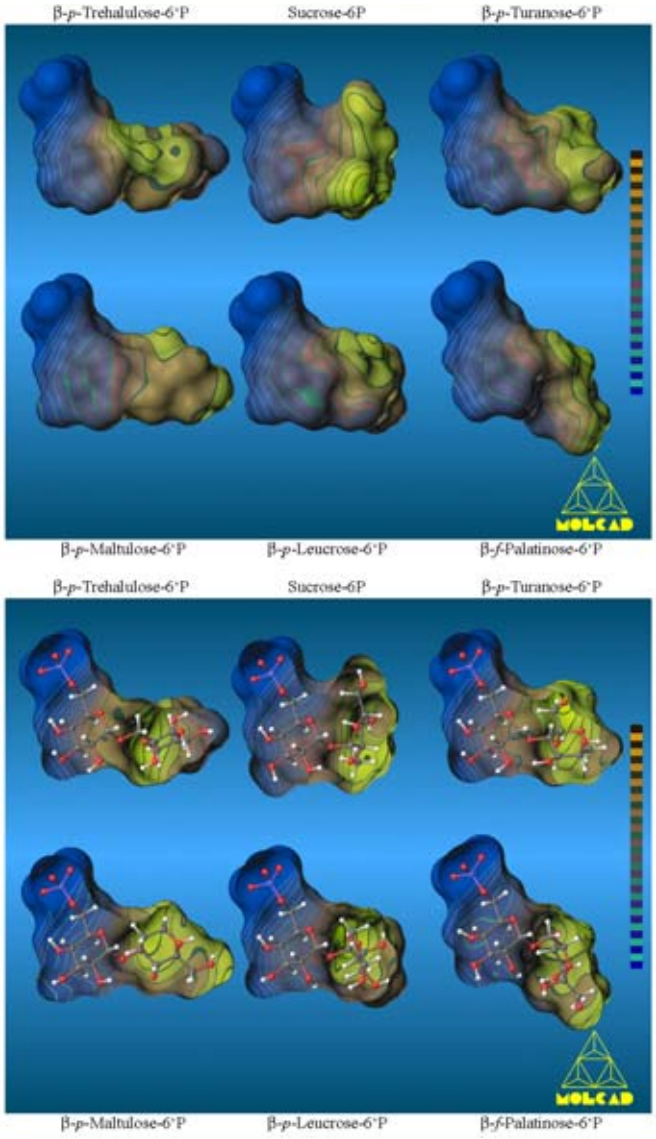


FIG. 10. **Molecular lipophilicity patterns (MLPs) of sucrose-6-P (*top center* entry) and its five isomeric 6′-phosphoglucosyl-fructoses in fully closed and front-side-opened form with *ball-and-stick* model inserts.** The relative hydrophobicity portraits were mapped in color-coded form onto their individual contact (solvent-accessible) surfaces with the colors ranging from *dark blue* (most hydrophilic areas) to *yellow-brown* (hydrophobic domains).

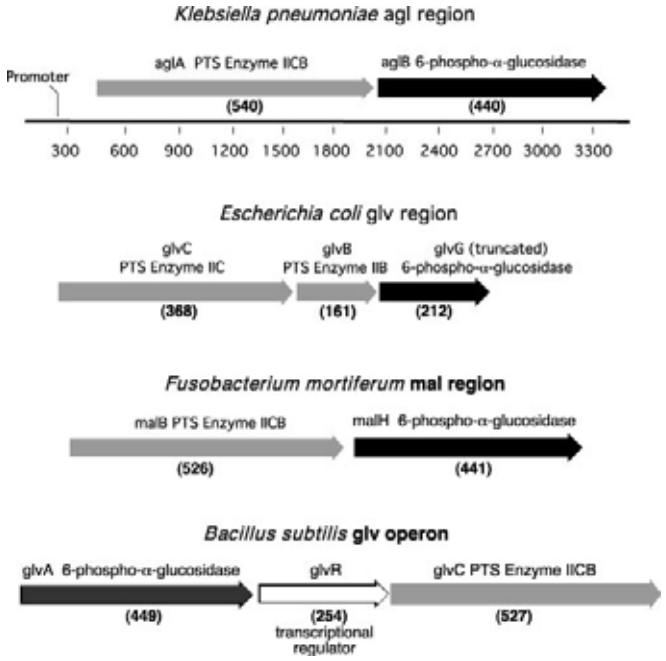


FIG. 11. **Comparison of the Agl region of *K. pneumoniae* (this paper; GenBank™ accession no. AF337811) with homologous regions of *E. coli* (Glv region (49)), *F. mortiferum* (Mal region; GenBank™ accession no. U81185 (44)), and *B. subtilis* (Glv operon; GenBank™ accession no. D50543 (60)).** Genetic elements are drawn to scale. Functionally equivalent genes are shown by the same types of *arrows*, and the *numbers in parentheses* indicate the number of residues encoded by the gene.

duced by growth of *K. pneumoniae* on the five glucosyl-fructoses, appears to be less specific and is tolerant of a considerable variation in both the structure and size of the *O*-linked aglycone. Indeed, the NAD^+ and metal ion-dependent phospho-α-glucosidase hydrolyzed not only the 6′-phosphoglucosyl-fructoses but also the phosphorylated derivatives of related α-linked disaccharides such as maltose-6′-P, isomaltose-6′-P, and maltitol-6-P. Remarkably, AglB was unable to hydrolyze sucrose-6-P. Explanations for enzyme specificity and substrate discrimination must reside in the molecular geometries and polarities of the individual disaccharide phosphates and/or in the three-dimensional structures of the two enzymes. Presently, only a structural model based on threading methods has been proposed for those enzymes (including sucrose-6-P hydrolase) that by sequence-based alignment are assigned to Family 32 of glycosyl hydrolases (57). Moreover, only a preliminary x-ray analysis has been reported for one enzyme member (GlvA from *B. subtilis*, (58)) of Family 4, to which AglB is assigned. Thus, we were led to probe the substrates with respect to structure, molecular shape, and polarity for clues to understanding the specificity of the two phosphoglucosyl hydrolases. From the markedly different molecular geometries of the phosphorylated disaccharides in solution (Figs. 8–10), one might reasonably assume that shape recognition (by the respective binding domains) may be an important determinant of enzyme specificity. Another and conceivably more significant contribution to substrate discrimination may originate from differences in the distribution of hydrophobic and hydrophilic regions over the contact surfaces of the disaccharide phosphates. In eliciting the sweetness response, for example, sucrose is believed to dock onto the taste bud receptor protein via its hydrophobic region (59), which, on the basis of the calculated MLP profiles, encompasses the entire outer surface side of the fructose moiety (51, 59). The same docking procedure is expected for sucrose-6-phosphate at the active site of sucrose-6-P hydrolase, inasmuch as the MLP profile of sucrose and its 6-phosphate (Fig. 10, *top center*) are essentially the same, *i.e.* a pronounced hydrophilic 6-phosphoglucosyl part (*blue areas*) facing a distinctly hydrophobic (*yellow*) fructose portion. Fig. 10 shows the MLPs of sucrose-6-P and its five isomeric 6′-phosphoglucosyl-fructoses in the fully closed (*upper portion*) and in the front-side-opened form with ball-and-stick model inserts (*lower portion*). The MLP patterns of the five 6′-phosphoglucosyl-fructoses, albeit having essentially identical hydrophilic (*blue*) glucose-6-P halves, clearly differ from sucrose-6-P with respect to the shape, intensity, and distribution of their hydrophobic (*yellow*) surface domains (Fig. 10). These may perhaps be the major factors that prevent docking of the isomeric phosphates at the sucrose-6-P binding site of sucrose-6-P hydrolase and, conversely, that preclude binding of sucrose-6-P to the active site of phospho-α-glucosidase.

Conclusion—This study and our earlier paper (12) are the first reports of bacterial growth on the five isomers of sucrose. However, genetic units similar to the Agl region of *K. pneu-*

moniae are present in the genomes of *B. subtilis*, *F. mortiferum*, and *E. coli* (Fig. 11). The phospho-α-glucosidase(s) of these species are clearly homologous (Fig. 5), and the PTS transporter (AglA) has extensive homology with GlvC of *B. subtilis,* MalB of *F. mortiferum*, and Glv(CB) of *E. coli*. The gene organization is similar in the three Gram-negative species, but for *B. subtilis* (Gram-positive*)* the gene order is reversed and a gene *glvR*, which encodes a regulatory protein, separates the phospho-α-glucosidase and PTS genes (60, 61). Our recent finding that *F. mortiferum* can also grow on the sucrose isomers[4] suggests that genes homologous to *aglA* and *aglB* may be prerequisites for bacterial growth on these compounds. Parenthetically, it may be noted that neither of these genes has been found during sequencing of the *S. mutans* genome, and these deficiencies may explain the inability of this organism to metabolize the sucrose isomers.

Acknowledgments—We express our thanks to Drs. Jack London and Edith C. Wolff for their encouragement, advice, and constructive criticisms. We also thank Drs. Nga Nguyen and Lewis Pannell for provision of microsequence and mass spectrometry data.

REFERENCES

1. Kundig, W., Ghosh, S., and Roseman, S. (1964) *Proc. Natl. Acad. Sci. U. S. A.* **52,** 1067–1074
2. Meadow, N. D., Fox, D. K., and Roseman, S. (1990) *Annu. Rev. Biochem.* **59,** 497–542
3. Postma, P. W., Lengeler, J. W., and Jacobson, G. R. (1993) *Microbiol. Rev.* **57,** 543–594
4. Reizer, J., Saier, M. H., Jr., Deutscher, J., Grenier, F., Thompson, J., and Hengstenberg, W. (1988) *Crit. Rev. Microbiol.* **15,** 297–338
5. Thompson, J. (1987) in *Sugar Transport and Metabolism in Gram-positive Bacteria* (Reizer, J., and Peterkofsky, A., eds) pp. 13–38, Ellis Horwood, Chichester, UK
6. Sprenger, G. A., and Lengeler, J. W. (1988) *J. Gen. Microbiol.* **134,** 1635–1644
7. Titgemeyer, F., Jahreis, K., Ebner, R., and Lengeler, J. W. (1996) *Mol. Gen. Genet.* **250,** 197–206
8. Kunst, F., Pascal, M., Lepesant, J-A., Walle, J., and Dedonder, R. (1974) *Eur. J. Biochem.* **42,** 611–620
9. Thompson, J., Nguygen, N. Y., Sackett, D. L., and Donkersloot, J. A. (1991) *J. Biol. Chem.* **266,** 14573–14579
10. IUPAC (1996) *IUPAC Recommendations for Nomenclature of Carbohydrates*, Academic Press, San Diego
11. Lichtenthaler, F. W., and Rönninger, R. (1990) *J. Chem. Soc. Perkin Trans.* **2,** 1489–1497
12. Thompson, J., Robrish, S. A., Pikis, A., Brust, A., and Lichtenthaler, F. W. (2001) *Carbohydr. Res.* **331,** 149–161
13. Loesche, W. J. (1986) *Microbiol. Rev.* **50,** 353–380
14. Newbrun, E. (1989) in *Cariology* (Newbrun, E., ed) 3rd Ed., Quintessence Publications Inc., Chicago
15. Takazoe, I., Frostell, G., Ohta, K., Topitsoglou, V., and Sasaki, N. (1985) *Swed. Dent. J.* **9,** 81–87
16. Ooshima, T., Izumitani, A., Takei, T., Fujiwara, T., and Sobue, S. (1990) *Caries Res.* **24,** 48–51
17. Moynihan, P. J. (1998) *J. Dentistry* **26,** 209–218
18. Schiweck, H., Munir, M., Rapp, K. M., Schneider, B., and Vogel, M. (1990) *Zuckerindustrie (Berlin)* **115,** 555–565
19. Schwengers, D. (1991) in *Carbohydrates As Organic Raw Materials* (Lichtenthaler, F. W., ed) pp. 183–195, VCH Publishers, New York
20. Thompson, J., Gentry-Weeks, C. R., Nguyen, N. Y., Folk, J. E., and Robrish, S. A. (1995) *J. Bacteriol.* **177,** 2505–2512
21. Sapico, V., Hanson, T. E., Walter, R. W., and Anderson, R. L. (1968) *J. Bacteriol.* **96,** 51–54
22. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem.* **4,** 187–217
23. Nilsson, L., and Karplus, M. (1986) *J. Comput. Chem.* **7,** 591–616
24. Reiling, S., Schlenkrich, M., and Brickmann, J. (1996) *J. Comput. Chem.* **17,** 450–468
25. Reiling, S., and Brickmann, J. (1995) *Macromol. Theory Simul.* **4,** 725–743
26. Brown, G. M., and Levy, H. A. (1973) *Acta Crystallogr. Sect. B Struct. Sci.* **29,** 790–797
27. Hanson, J. C., Sieker, L. C., and Jensen, L. H. (1973) *Acta Crystallogr. Sect. B Struct. Sci.* **29,** 797–808
28. Kanters, J. A., Gaykema, W. P. J., and Roelofsen, G. (1978) *Acta Crystallogr. Sect. B Struct. Sci.* **34,** 1873–1881
29. Thiem, J., Kleeberg, M., and Klaska, K.-H. (1989) *Carbohydr. Res.* **189,** 65–77
30. Dreissig, V. W., and Luger, P. (1973) *Acta Crystallogr. Sect. B Struct. Sci.* **29,** 514–521
31. Jeffrey, G. A., and Nanni, R. (1985) *Carbohydr. Res.* **137,** 21–30
32. Quigley, G. J., Sarko, A., and Marchessault, R. H. (1970) *J. Am. Chem. Soc.* **92,** 5834–5839
33. Gress, M. E., and Jeffrey, G. A. (1977) *Acta Crystallogr. Sect. B Struct. Sci.* **33,** 2490–2495
34. Ohno, S., Hirao, M., and Kido, M. (1982) *Carbohydr. Res.* **108,** 163–171
35. Schouten, A., Kanters, J. A., Kroon, J., Looten, P., Duflot, P., and Mathlouthi, M. (1999) *Carbohydr. Res.* **322,** 298–302
36. Cambridge Crystallographic Data Center (2001) *Cambridge Structural Database*, Version 5.21, Cambridge Crystallographic Data Center, Cambridge, UK
37. Allen, F. H., Bellard, S. A., Brice, M. D., Cartwright, B. A., Doubleday, A., Higgs, H., Hummelink, T., Hummelink-Peters, B. G., Kennard, O., Motherwell, W. D. S., Rodgers, J. R., and Watson, D. G. (1979) *Acta Crystallogr. Sect. B Struct. Sci.* **35,** 2331–2339
38. Connolly, M. L. (1983) *Science* **221,** 709–713
39. Heiden, W., Moeckel, G., and Brickmann, J. (1993) *J. Comput. Aided Mol. Des.* **7,** 503–514
40. Teschner, M., Henn, C., Vollhardt, H., Reiling, S., and Brickmann, J. (1994) *J. Mol. Graph.* **12,** 98–105
41. Brickmann, J. (1996) *MOLCAD-Molecular Computer Aided Design*, Darmstadt University of Technology, Germany
42. Waldherr-Teschner, M., Goetze, T., Heiden, W., Knoblauch, M., Vollhardt, H., and Brickmann, J. (1992) in *Advances in Scientific Visualization* (Post, F. H., and Hin, A. J. S., eds) pp. 58–67, Springer, Heidelberg, Germany
43. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipmann, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
44. Bouma, C. L., Reizer, J., Reizer, A., Robrish, S. A., and Thompson, J. (1997) *J. Bacteriol.* **179,** 4129–4137
45. Thompson, J., Pikis, A., Ruvinov, S. B., Henrissat, B., Yamamoto, H., and Sekiguchi, J. (1998) *J. Biol. Chem.* **273,** 27347–25356
46. Lengeler, J. W., Jahreis, K., and Wehmeier, U. F. (1994) *Biochim. Biophys. Acta* **1188,** 1–28
47. Henrissat, B. (1991) *Biochem. J.* **280,** 309–316
48. Swiss Institute of Bioinformatics (2000) *SWISS-PROT, Protein Sequence Data Bank*, Release 39.0, SIB, Geneva
49. Reizer, J., Michotey, V., Reizer, A., and Saier, M. H., Jr. (1994) *Protein Sci.* **3,** 440–450
50. Lichtenthaler, F. W., Immel, S., and Kreis, U. (1991) *Starch/Stärke* **43,** 121–132
51. Lichtenthaler, F. W., and Immel, S. (1995) *Int. Sugar J.* **97,** 12–22
52. Immel, S., and Lichtenthaler, F. W. (1995) *Liebigs Ann.* 1925–1937
53. Thompson, J., Ruvinov, S. B., Freedberg, D. I., and Hall, B. G. (1999) *J. Bacteriol.* **181,** 7339–7345
54. Raasch, C., Streit, W., Schanzer, J., Bibel, M., Gosslar, U., and Liebl, W. (2000) *Extremophiles* **4,** 189–200
55. Campbell, J. A., Davies, G. J., Bulone, V., and Henrissat, B. (1997) *Biochem. J.* **326,** 929–939
56. Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A., and Negishi, M. (2000) *J. Biol. Chem.* **275,** 34580–34585
57. Pons, T., Olmea, O., Chinea, G., Beldarrain, A., Márquez, G., Acosta, N., Rodríguez, L., and Valencia, A. (1998) *Proteins* **33,** 383–395
58. Varrot, A., Yamamoto, H., Sekiguchi, J., Thompson, J., and Davies, G. J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 1212–1214
59. Lichtenthaler, F. W., and Immel, S. (1993) in *Correlations of Hydrophobicity Potential Profiles of Sucrose, Sucralose and Fructose with AH-B-X Assignment in Sweet Taste Chemoreception* (Mathlouthi, M., Kanters, J. A., and Birch, G. G., eds) pp. 21–53, Elsevier Science, New York
60. Yamamoto, H., Uchiyama, S., Fajar, A. N., Ogasawara, N., and Sekiguchi, J. (1996) *Microbiology-UK* **142,** 1417–1421
61. Yamamoto, H., Serizawa, M., Thompson, J., and Sekiguchi, J. (2001) *J. Bacteriol.* **183,** 5110–5121

[4] A. Pikis, S. Immel, S. A. Robrish, and J. Thompson, unpublished data.

Microbiology (2002), **148**, 843–852 Printed in Great Britain

Metabolism of sucrose and its five isomers by *Fusobacterium mortiferum*

Andreas Pikis,[1,2] Stefan Immel,[3] Stanley A. Robrish[1] and John Thompson[1]

Author for correspondence: John Thompson. Tel: +1 301 496 4083. Fax: +1 301 402 0396. e-mail: jthompson@dir.nidcr.nih.gov

[1] Microbial Biochemistry and Genetics Unit, Oral Infection and Immunity Branch, National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, MD 20892-4350, USA

[2] Department of Infectious Diseases, Children's National Medical Center, Washington DC 20010-2970, USA

[3] Institut für Organische Chemie, Technische Universität Darmstadt, D-64287 Darmstadt, Germany

Fusobacterium mortiferum **utilizes sucrose [glucose-fructose in α(1→2) linkage] and its five isomeric α-D-glucosyl-D-fructoses as energy sources for growth. Sucrose-grown cells are induced for both sucrose-6-phosphate hydrolase (S6PH) and fructokinase (FK), but the two enzymes are not expressed above constitutive levels during growth on the isomeric compounds. Extracts of cells grown previously on the sucrose isomers trehalulose α(1→1), turanose α(1→3), maltulose α(1→4), leucrose α(1→5) and palatinose α(1→6) contained high levels of an NAD^+ plus metal-dependent phospho-α-glucosidase (MalH). The latter enzyme was not induced during growth on sucrose. MalH catalysed the hydrolysis of the 6′-phosphorylated derivatives of the five isomers to yield glucose 6-phosphate and fructose, but sucrose 6-phosphate itself was not a substrate. Unexpectedly, MalH hydrolysed both α- and β-linked stereomers of the chromogenic analogue *p*-nitrophenyl glucoside 6-phosphate. The gene *malH* is adjacent to *malB* and *malR*, which encode an EII(CB) component of the phospho*enol*pyruvate-dependent sugar:phosphotransferase system and a putative regulatory protein, respectively. The authors suggest that for *F. mortiferum*, the products of *malB* and *malH* catalyse the phosphorylative translocation and intracellular hydrolysis of the five isomers of sucrose and of related α-linked glucosides. Genes homologous to *malB* and *malH* are present in both *Klebsiella pneumoniae* and the enterohaemorrhagic strain *Escherichia coli* O157:H7. Both these organisms grew well on sucrose, but only *K. pneumoniae* exhibited growth on the isomeric compounds.**

Keywords: phospho-α-glucosidase, sucrose isomers, sucrose-6-phosphate hydrolase, *Klebsiella pneumoniae*, *Escherichia coli* O157:H7

INTRODUCTION

Many bacterial species, including *Klebsiella pneumoniae* (Sprenger & Lengeler, 1988; Titgemeyer *et al*., 1996), *Bacillus subtilis* (Fouet *et al*., 1987), *Lactococcus lactis* (Thompson & Chassy, 1981; Thompson *et al*., 1991; Rauch & deVos, 1992), *Fusobacterium mortiferum* (Thompson *et al*., 1992), *Escherichia coli* (Schmid *et al*., 1988) and *Clostridium beijerinckii* (Tangney *et al*., 1998; Reid *et al*., 1999) translocate sucrose simultaneously with phosphorylation at C-6 of the glucosyl moiety via the phospho*enol*pyruvate-dependent sucrose:phosphotransferase system (PEP:PTS) (Meadow *et al*., 1990; Postma *et al*., 1993). Sucrose 6-phosphate is hydrolysed intracellularly by sucrose-6-phosphate hydrolase (S6PH) to yield glucose 6-phosphate and fructose, which are further metabolized via the glycolytic pathway. The multi-component sucrose–PEP:PTS and S6PH are also expressed by oral streptococci, including *Streptococcus mutans* (St Martin & Wittenberger, 1979; Slee & Tanzer, 1979) and *Streptococcus sobrinus* (Chen & LeBlanc; 1992) and dietary sucrose is fermented primarily to lactic acid. By its demineralizing action upon tooth enamel, this organic acid initiates or contributes to the aetiology of dental caries (Loesche, 1986; Van Houte, 1994).

The linkage between the two component sugars of sucrose, i.e. D-glucose and D-fructose, can be modified to

Abbreviations: FK, fructokinase; G6P, glucose 6-phosphate; G6PDH, glucose-6-phosphate dehydrogenase; HK, hexokinase; 4-MU-α-G6P, 4-methylumbelliferyl α-D-glucopyranoside-6-phosphate; PEP:PTS, phospho-*enol*pyruvate-dependent sucrose:phosphotransferase system; pNP-α-G6P, *p*-nitrophenyl α-D-glucopyranoside 6-phosphate; pNP-β-G6P, *p*-nitrophenyl β-D-glucopyranoside 6-phosphate; S6PH, sucrose-6-phosphate hydrolase.

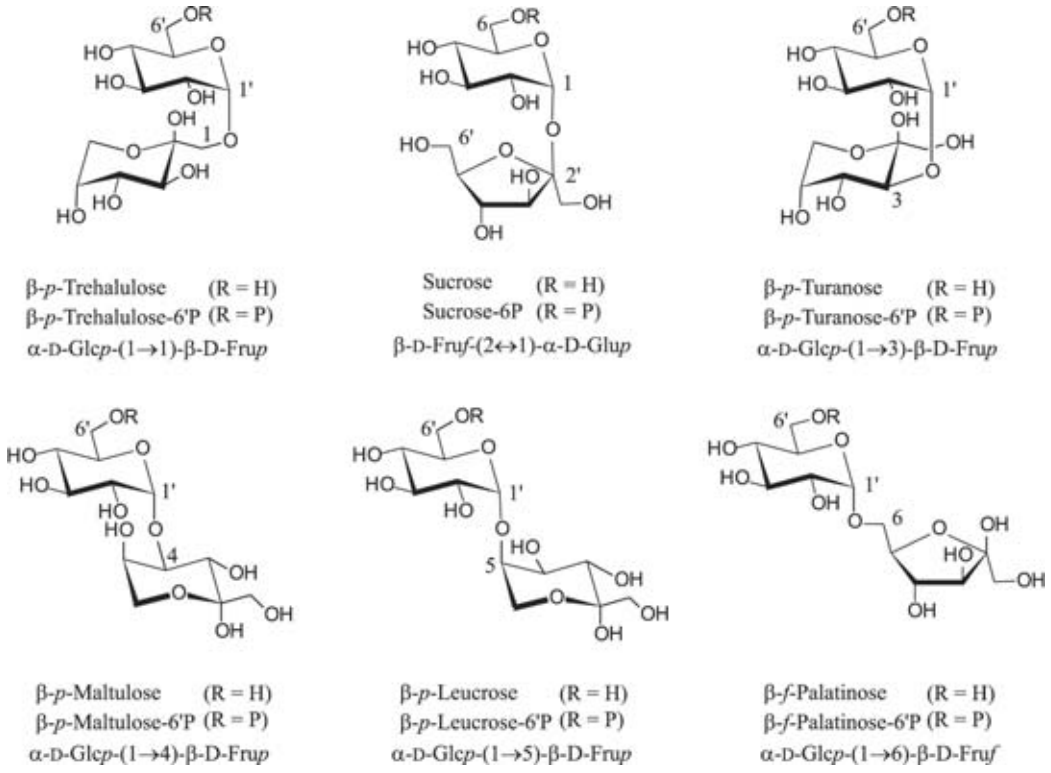


Fig. 1. Molecular formulae of sucrose and its five isomeric α-D-glucosyl-D-fructoses, in both free and phosphorylated forms.

yield five isomeric compounds (Fig. 1) that trivially are designated trehalulose, turanose, maltulose, leucrose and palatinose (Lichtenthaler & Rönninger, 1990; Lichtenthaler *et al*., 1991; Immel & Lichtenthaler, 1995). Until recently (Thompson *et al*., 2001a), there were no reports of the bacterial utilization of the five sucrose isomers. Indeed, the inability of mutans streptococci to metabolize these comparatively sweet disaccharides suggests their use as non-cariogenic substitutes for dietary sucrose (Ooshima *et al*., 1983, 1991; Ziesenitz *et al*., 1989; Minami *et al*., 1990; Peltroche-Llacsahuanga *et al*., 2001). In light of these reports, we were surprised to discover that *K. pneumoniae* readily utilized all five α-D-glucosyl-D-fructoses as energy sources for growth (Thompson *et al*., 2001a, b), and that the enzymes encoded by the sucrose (*scr*) operon (Titgemeyer *et al*., 1996) did not participate in dissimilation of these compounds. Remarkably, the sucrose isomers and structurally related α-glucosides (including maltose, isomaltose, maltitol and methyl α-D-glucoside) are first translocated by an α-glucoside-specific EII(CB) transport protein, and the accumulated 6-phospho-α-D-glucosides are hydrolysed by a metal-requiring, NAD^+-dependent phosphoglucosyl hydrolase belonging to family 4 of the glycosylhydrolase superfamily (Henrissat, 1991). In *K. pneumoniae*, the genes for the EII(CB) transport protein (*aglA*) and the phospho-α-glucosidase (*aglB*) lie adjacent, and are chromosomally encoded (Thompson *et al*., 2001b).

Fusobacteria are Gram-negative anaerobic rods that are usually described as weakly or asaccharolytic, and most species, including *Fusobacterium nucleatum*, use amino acids as fermentable energy sources (Robrish *et al*., 1987; Robrish & Thompson, 1990). The products of metabolism (acetic, butyric and propionic acids) may penetrate periodontal tissue, thereby contributing to the aetiology of gingivitis and periodontal disease. In contrast to other species, *F. mortiferum* ferments an extraordinarily wide variety of carbohydrates (Robrish *et al*., 1991). Previously, in studies of maltose metabolism in *F. mortiferum*, we cloned and expressed a gene (*malH*) whose deduced sequence exhibits ~75% residue identity with the phospho-α-glucosidase of *K. pneumoniae* (Thompson *et al*., 1995; Bouma *et al*., 1997). In the present report, we show that the gene adjacent to *malH* (designated *malB*) also encodes a putative EII(CB) protein that is ~60% identical with AglA of *K. pneumoniae*. Coincident with our studies, publication of the complete genome sequence of enterohaemorrhagic *E. coli* O157:H7 (Perna *et al*., 2001) also revealed two adjacent genes with extensive homology to those found in *K. pneumoniae*. It was of interest, therefore, to determine whether possession of these genetic elements would also permit growth of *F. mortiferum* and *E. coli* O157:H7 on the five isomers of sucrose. Our findings are summarized in this communication. Additionally, we describe the purification, and some unexpected properties, of the phospho-α-

glucosidase (MalH) that catalyses the hydrolysis of phosphorylated sucrose isomers in *F. mortiferum*.

METHODS

Materials and reagents. Sucrose isomers were obtained from the following sources: trehalulose, Südzucker, Mannheim/Ochsenfurt; turanose, Pfanstiehl Laboratories; maltulose, TCI America; leucrose, Fluka; palatinose, Wako Chemicals. Other sugars and glucosides were purchased from Sigma and Pfanstiehl Laboratories. Phosphorylated derivatives were biosynthesized via the PEP:PTS of permeabilized (palatinose-grown) cells of *K. pneumoniae* and were purified by Ba^{2+}/ethanol precipitation, ion-exchange and paper chromatography (Thompson *et al.*, 2001a). The chromogenic and fluorogenic substrates *p*-nitrophenyl α-D-glucopyranoside 6-phosphate (pNP-α-G6P), *p*-nitrophenyl β-D-glucopyranoside 6-phosphate (pNP-β-G6P) and 4-methylumbelliferyl α-D-glucopyranoside-6-phosphate (4-MU-α-G6P) were prepared by selective phosphorylation (at C6-OH) of the parent glucosides with phosphorus oxychloride in trimethyl phosphate containing small proportions of water (Thompson *et al.*, 1995). Glucose-6-phosphate dehydrogenase (G6PDH, EC 1.1.1.49) and hexokinase (HK, EC 2.7.1.1) were purchased from Boehringer Mannheim, and Ultrogel AcA-44 and Tris-Acryl M-DEAE from Sigma.

Bacterial strains and culture media. *K. pneumoniae* ATCC 23357, *E. coli* O157:H7 (EDL 933) and *F. mortiferum* ATCC 25557 were obtained from the American Type Culture Collection. *K. pneumoniae* and *E. coli* O157:H7 were grown in a medium of the following composition (per litre): Na_2HPO_4, 7·1 g; KH_2PO_4, 1·5 g; $(NH_4)_2SO_4$, 3 g; $MgSO_4.7H_2O$, 0·1 g; $FeSO_4.7H_2O$, 5 mg. Filter-sterilized sugars were added to autoclaved media (pH 7·4) to a final concentration of 4 g per litre. Cells of *K. pneumoniae* were grown in standing cultures, but *E. coli* O157:H7 was grown with vigorous aeration on a rotary shaker (~ 250 r.p.m.). *E. coli* PEP43(pCB4.11) was grown with aeration in Luria–Bertani (LB) medium supplemented with 50 μg kanamycin ml^{-1}. *F. mortiferum* was grown anaerobically (GasPak, BBL) in a medium comprising (per litre): Tryptone (Difco), 17 g; Protease Peptone (Difco), 3 g; Na_2HPO_4, 2·5 g; NaCl, 5 g; final pH 7·3.

DNA sequence analysis. Automated DNA sequencing incorporating Big Dye terminators was used to sequence *malB*, and an adjacent upstream gene (*malR*), directly from genomic DNA of *F. mortiferum*. From data previously reported by Bouma *et al.* (1997), a reverse primer 1R1 (5′-AACTCT-CTCTAACTTGTGGTACTGAAAGTC-3′) was designed to obtain initial sequence information. Subsequent data were obtained by the primer synthesis and the chromosomal 'walking' technique. PCR primers were designed for sequencing of the second strand, and for amplification (from genomic DNA) of the fragment encoding the two genes by use of *Taq* DNA polymerase. The amplicon was cloned into the TOPO TA cloning vector (Invitrogen). All sequencing was performed by BioServe Biotechnologies (Laurel, MD, USA), and the MacVector 7.0 sequence analysis package (Genetics Computer Group, Madison, WI, USA) was used to assemble and analyse the data.

Metabolism of sugars by washed cells of *F. mortiferum*. To maintain anaerobic conditions, centrifuge tubes were flushed with a gas mixture (5 % CO_2, 5 % H_2, 90 % N_2, by vol.) prior to harvesting of the sucrose-grown cells (5000 ***g*** for 10 min at 5 °C). The supernatant fluid was discarded, and the cell pellet was resuspended as quickly as possible in 30 ml anaerobically prepared wash solution [50 mM potassium phosphate buffer (pH 7) containing 1 mM $MgCl_2$]. After centrifugation, the washed cells were resuspended in 5 ml wash buffer, and the mixture was maintained at 0 °C under anaerobic conditions until required. For studies of disaccharide utilization, the washed cells (equivalent to 30 mg total cell protein) were added to 10 ml of wash buffer containing the desired sugar (sucrose or isomer) at a final concentration of 10 mM. The cell suspensions were incubated at 37 °C in 100 ml serum bottles filled with anaerobic gas and, at intervals, 1 ml samples were withdrawn by insertion of a gas-flushed tuberculin syringe through the butyl rubber cap. The cells were removed by filtration through Millex-GS filter units (0·22 μm pore size; Millipore) and filtrates were collected. Samples were heated in 1 M HCl for 1 h at 100 °C, cooled, and neutralized with 1 M KOH. Glucose (equivalent to disaccharide remaining) was determined by the ATP–G6PDH/HK–$NADP^+$ coupled enzyme assay.

Preparation and analysis of *F. mortiferum* extracts. Cells of *F. mortiferum* grown on the different sugars were harvested from 400 ml anaerobic culture. The cell pellets (~ 1–2 g wet weight) were resuspended with 3 vols 25 mM Tris/HCl buffer (pH 7·5) containing 0·1 mM NAD^+ and 1 mM $MnCl_2$ (designated TNM buffer). The cells were disrupted by sonication at 0 °C (6 × 15 s bursts in a Branson instrument, model 185), and centrifuged at 14000 r.p.m. for 20 min at 5 °C in an Eppendorf bench-top instrument. The clarified supernatants were assayed for S6PH, FK and phospho-α-glucosidase activities.

Enzyme assays. The activities of S6PH, FK and phospho-α-glucosidase (with disaccharide phosphate substrates) were determined from the formation of glucose, fructose 6-phosphate and G6P, respectively, in the appropriate reaction mixture. Production of the three metabolites was coupled to the enzymic reduction of $NADP^+$ (measured as A_{340}), and rates were determined in a Beckman DU 640 recording spectrophotometer. In all calculations, a molar absorption coefficient (ε) of 6220 M^{-1} cm^{-1} was assumed for NADPH.

S6PH. This enzyme catalyses the hydrolysis of both sucrose 6-phosphate (to G6P and fructose) and sucrose (to glucose and fructose), albeit with significantly different K_m for the two compounds (0·1 mM and ~ 100 mM, respectively; see Thompson *et al.*, 1992). Because of the limited supply of sucrose 6-phosphate, sucrose was used as substrate for the spectrophotometric assay of S6PH in cell extracts. The standard 1 ml assay contained: 0·1 M potassium phosphate buffer (pH 7·2); 50 mM sucrose; 5 mM ATP; 10 mM $MgCl_2$; 1 mM $NADP^+$, ~ 3 U each of G6PDH/HK, and cell extract.

FK. Activity was determined in a similar mixture to that used for S6PH, containing 10 mM fructose as substrate, 3 U G6PDH and 5 U phosphoglucose isomerase.

Phospho-α-glucosidase (MalH). Activity was determined by two methods in which either chromogenic analogues or phosphorylated disaccharides served as substrates. Cofactors NAD^+ and Mn^{2+} were included in both reaction mixtures. Throughout the purification of MalH, enzyme activity was determined in a discontinuous assay with pNP-α-G6P and pNP-β-G6P as substrates. The 2 ml reaction mixture (at 37 °C) contained: 50 mM Tris/HCl buffer (pH 7·5); 0·5 mM NAD^+; 1 mM $MnCl_2$; and 0·5 mM of the chromogenic substrate. After enzyme addition, samples (0·25 ml) were removed at intervals throughout a 3 min period of incubation, and were immediately added to 0·75 ml 0·5 M Na_2CO_3 solution containing 0·1 M EDTA to stop the reaction. The A_{400} was

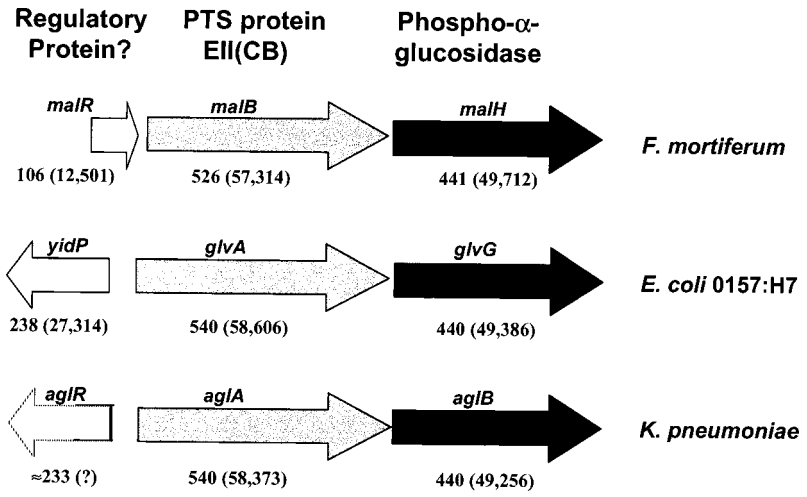


Fig. 2. Structural organization of the putative α-glucoside operons of: *F. mortiferum* (GenBank accession number U81185); *E. coli* O157:H7 (*yidP*, GenBank AAG58886; *glvA*, GenBank AAG58885; *glvG*, GenBank AAG58884) and *K. pneumoniae* (GenBank AF337811). The numbers below the arrows are predicted amino acid residues, with calculated molecular masses (Da) of the encoded polypeptides in parentheses. Dotted lines indicate incomplete sequence for *aglR* of *K. pneumoniae*.

measured, and the amount of pNP formed (substrate hydrolysed) was calculated from the molar absorption coefficient of the yellow *p*-nitrophenolate anion, $\varepsilon = 18300\ M^{-1}\ cm^{-1}$. A continuous $NADP^+$-coupled assay was used to measure G6P formed by MalH-catalysed hydrolysis of phosphorylated disaccharides. The assay mixture contained in 1 ml: 0·1 M HEPES buffer (pH 7·5); 1 mM $MgCl_2$; 1 mM $MnCl_2$; 1 mM NAD^+; 1 mM $NADP^+$; 2 mM disaccharide phosphate, 3 U G6PDH and purified MalH (35 μg protein).

Purification of MalH from *E. coli* PEP43(pCB4.11). A 2·2 kb *Sau*3AI chromosomal DNA fragment of *F. mortiferum* that includes the *malH* gene and its promoter has previously been cloned and the enzyme has been expressed from plasmid pCB4.11 (Bouma *et al*., 1997). This plasmid was transferred by electroporation to *E. coli* PEP43 Δ*cel* Δ(*bgl–pho*) *leu metA* or *B his rpsL lacZ*Δ4680 *lacY*$^+$ *arbT*$^+$ Tn*10*::*bglA* dTn*10*$_{cam}$::*ebgA*5100 *ebgR*$^+$ L532 (B. G. Hall, Biology Department, University of Rochester, NY, USA, laboratory collection). *E. coli* PEP43 expresses no phospho-*β*-glucosidases because the *cel* and *bglGFB* operons have been deleted, *bglA* is disrupted by Tn*10* and the *asc* operon is cryptic.

Cells of *E. coli* PEP43(pCB4.11) (~ 25 g wet weight) were resuspended with 40 ml TNM buffer, and the organisms were disrupted by 2 × 1·5 min sonication with a Branson model 350 instrument. The preparation was clarified by ultracentrifugation (180000 ***g*** for 2 h at 5 °C), and the supernatant fluid was dialysed against 4 litres of TNM buffer. The dialysed material was transferred (~ 0·6 ml min^{-1}) to a column of TrisAcryl M-DEAE (2·6 × 10 cm) that had been equilibrated with TNM buffer. The column was washed to elute non-adsorbed material, and then MalH activity was eluted with 500 ml of a linear, increasing concentration gradient of NaCl (0–150 mM) in TNM buffer. Fractions of 5 ml were collected, and those containing highest MalH activity (54–65 inclusive) were pooled and concentrated in an Amicon pressure cell to ~ 3 ml. The concentrated sample was transferred (0·15 ml min^{-1}) to an Ultrogel AcA-44 gel filtration column (1·6 × 94 cm; linear fractionation range, 10–130 kDa) previously equilibrated with TNM buffer containing 0·1 M NaCl. Fractions of 2 ml were collected, and tetrameric MalH (~ 200 kDa) was eluted at the void volume of the column. Fractions that contained a single protein by SDS-PAGE (47–50, inclusive) were pooled, and concentrated to yield ~ 5 mg purified MalH [specific activity 2·9 μmol pNP-α-G6P hydrolysed min^{-1} (mg protein)$^{-1}$].

Analytical methods. Protein concentrations were determined by the BCA protein assay (Pierce). The Novex X-Cell system was used for both native (nonreducing) and SDS-PAGE. For SDS-PAGE experiments, precast NuPage (4–12 %) Bistris gels and MES-SDS running buffer (pH 7·3) were used with Novex Mark 12 protein size standards. Proteins were stained with Coomassie brilliant blue R-250. Electrophoresis of cell extracts under nonreducing conditions was carried out at 10 °C in Tris/glycine (4–20 %) precast gels from Novex, with Tris/glycine (pH 8·3) supplemented with 1 mM $MnCl_2$ and 0·1 mM NAD^+ as the running buffer. For detection of phospho-α-glucosidase activity, the gel was immersed in 30 ml of a solution that contained 25 mM Tris/HCl buffer (pH 7·5); 1 mM $MnCl_2$; 0·1 mM NAD^+ and 0·1 mM 4MU-α-G6P. After ~ 5 min incubation, the gel was photographed under long-wave UV light with Ektopan Kodak film (2 min exposure with a green filter). For Western blot analysis, proteins in the cell extracts together with pre-stained markers (SeeBlue from Novex), were first separated by SDS-PAGE and then transferred to a nitrocellulose membrane. Immunodetection of phospho-α-glucosidase was performed with polyclonal antibody to MalH from *F. mortiferum* as described previously (Thompson *et al*., 1995). Molecular dynamics simulations and procedures for the determination of solvent-accessible surfaces have been described previously (Immel & Lichtenthaler, 1995; Thompson *et al*., 2001b).

RESULTS

Gene organization in *F. mortiferum*, *K. pneumoniae* and *E. coli* O157:H7

The genes that constitute the putative α-glucoside operons, and their organization in the three bacterial species, are shown in Fig. 2. In *K. pneumoniae*, adjacent chromosomal genes *aglA* and *aglB* encode, respectively, an EII(CB) transport protein and phospho-α-glucosidase. These proteins promote the phosphorylative translocation and hydrolysis of sucrose isomers by this organism (Thompson *et al*., 2001a, b). The partial

Table 1. Enzyme activities in extracts prepared from cells of *F. mortiferum* grown previously on various carbohydrates

Growth sugar	Enzyme specific activity†		
	Phospho-α-glucosidase‡	FK§	S6PH‖
Trehalulose*	11·6	2·6	2·2
Sucrose	ND	26·3	63·4
Turanose*	3·5	0·7	1·9
Maltulose*	3·4	1·6	1·6
Leucrose*	12·7	2·4	6·5
Palatinose*	3·2	1·2	4·0
Glucose	ND	2·1	ND
Fructose	ND	1·0	3·6
Maltose	8·5	3·1	2·8
Methyl α-D-glucoside	23·9	ND	2·8

* Sucrose isomers.

† The same cell extracts were used for the assay of the three enzyme activities; values are means of two separate assays. ND, No detectable activity.

‡ nmol pNP-α-G6P hydrolysed min^{-1} (mg protein)$^{-1}$.

§ nmol fructose phosphorylated min^{-1} (mg protein)$^{-1}$.

‖ nmol sucrose hydrolysed min^{-1} (mg protein)$^{-1}$.

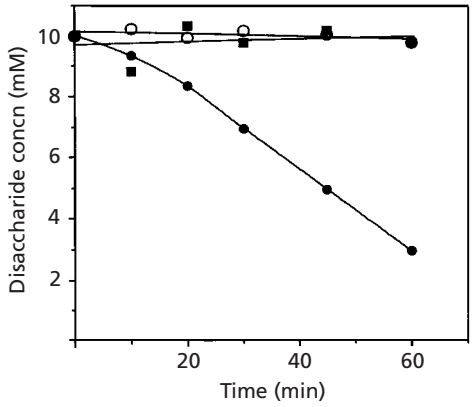


Fig. 3. Studies of the metabolism of sucrose and two of its isomers by sucrose-grown cells of *F. mortiferum*. Washed cells were resuspended, under anaerobic conditions, in buffered solution containing the disaccharides at an initial concentration of 10 mM. Samples were removed at intervals and residual disaccharide was determined by the enzymic assay of glucose produced by acid hydrolysis. ●, Sucrose; ○, turanose; ■, palatinose.

Table 2. Expression of S6PH and FK during growth of *K. pneumoniae* on different sugars

Growth sugar	Enzyme specific activity	
	FK†	S6PH‡
Trehalulose*	69	289
Sucrose	103	247
Turanose*	64	342
Maltulose*	54	170
Leucrose*	51	184
Palatinose*	87	289
Maltose	15	29
Trehalose	3	ND
Melibiose	9	18
Cellobiose	2	6
Maltitol	1	3
Glucose	ND	2
Methyl α-D-glucoside	1	ND
Galactose	8	6

ND, No detectable activity.

* Sucrose isomers.

† nmol fructose phosphorylated min^{-1} (mg protein)$^{-1}$.

‡ nmol sucrose hydrolysed min^{-1} (mg protein)$^{-1}$.

sequence of the regulatory gene, *aglR*, was compiled from our own work, together with data obtained from the Washington University (St Louis) sequencing project of the *K. pneumoniae* genome (http://genome.wustl.edu). The recently published genome sequence of enterohaemorrhagic *E. coli* O157:H7 (Perna *et al.*, 2001) revealed the same organization of the three genes (designated *yidP*, *glvA* and *glvG*) as described for *K. pneumoniae*. The amino acid sequence deduced from *yidP* of *E. coli* O157:H7 predicts a polypeptide of 238 residues that exhibits 91% overall identity with the 233 residues deduced by translation of the incomplete gene *aglR* of *K. pneumoniae*. At their N-termini, the products of *aglR* and *yidP* contain a helix–turn–helix (HTH)

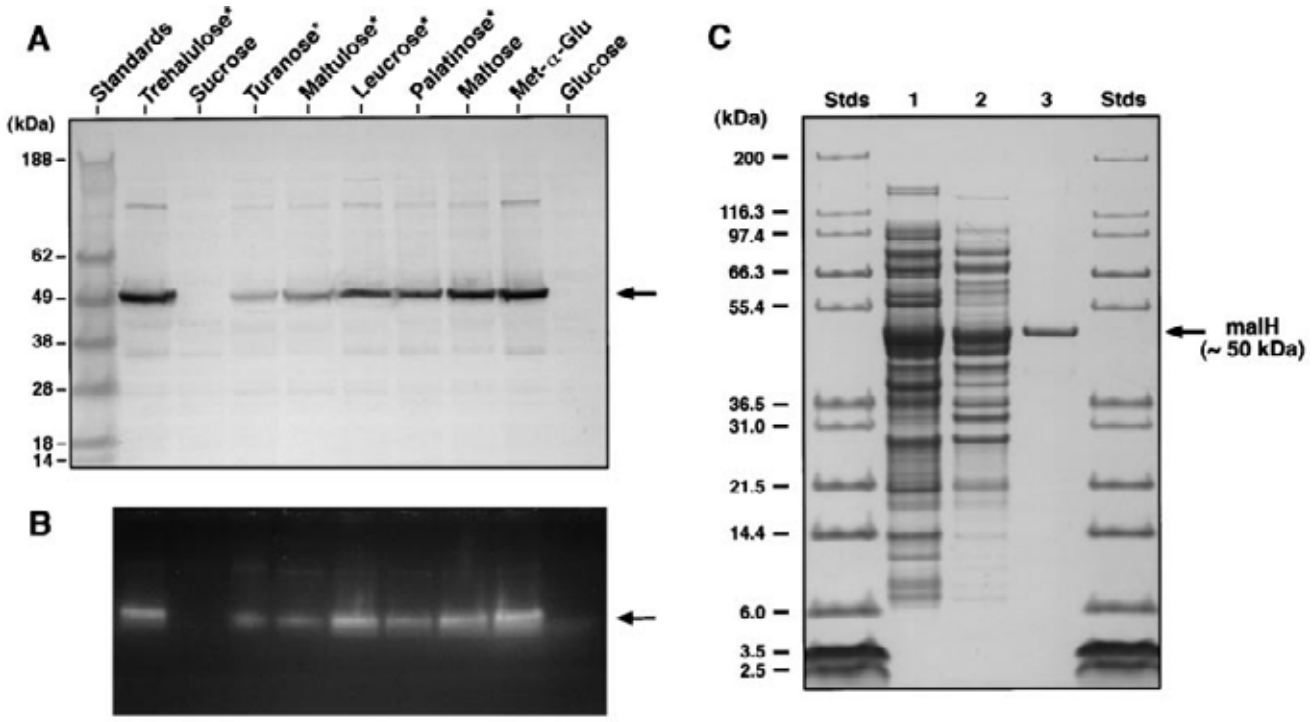


Fig. 4. Demonstration by PAGE of the expression, *in situ* activity, and purification of phospho-α-glucosidase (MalH) from *F. mortiferum*. (A) Immunodetection of MalH expression during growth of *F. mortiferum* on different sugars, by Western blotting using polyclonal antibody to MalH. Note the absence of immunoreactive polypeptide (~ 50 kDa, arrow) in the extract from sucrose-grown cells. (B) Zymogram demonstration of MalH activity in cell extracts by hydrolysis (arrow) of the fluorogenic substrate 4MUαG6P. Again, note the absence of fluorescence in the lane containing the extract from sucrose-grown cells. (C) Purification and determination of the molecular mass of MalH. Samples from each of the three stages of purification were denatured, and polypeptides were resolved by SDS-PAGE: lane 1, dialysed high-speed supernatant; lane 2, TrisAcryl M-DEAE; lane 3, purified MalH (molecular mass ~ 50 kDa) obtained by AcA-44 gel filtration chromatography. The asterisks in panel A indicate sucrose isomers.

motif that assigns the two proteins to the GntR family of transcriptional regulators. The complete sequence of the phospho-α-glucosidase gene (*malH*) of *F. mortiferum*, together with a partial sequence for the gene (*malB*), were described in an earlier report (Bouma *et al.*, 1997). The entire sequences for *malB*, and the upstream gene (*malR*), have now been obtained by chromosome 'walking' (GenBank accession no. U81185). Translation of *malR* predicts a 106-residue polypeptide that shows extensive homology with the 13 members of the UPF0087 family of regulatory proteins.

Sequence alignment of phospho-α-glucosidase and EII(CB) proteins

Alignment of the amino acid sequences predicted for the phospho-α-glucosidase(s) and EIIs reveals a high degree of similarity among these proteins (data not shown). MalH from *F. mortiferum* exhibits ~75% residue identity with the phosphoglucosyl hydrolase(s) from *E. coli* O157:H7 and *K. pneumoniae*. The EII(CB) transport protein of *F. mortiferum* (MalB) shows ~60% amino acid identity throughout its length with GlvA and AglA in *E. coli* O157:H7 and *K. pneumoniae*, respectively. For the two enteric species, phospho-α-glucosidase and EII(CB) proteins exhibit overall identities of 89% and 81%, respectively. By sequence-based alignment (Henrissat, 1991) and signature pattern (P-X-[SA]-X-[LIVMFY](2)-[QN]-X(2)-N-P-X(4)-[TA]-X(9,10)-[KRD]-X-[LIV]-[GN]-X-C), the three phospho-α-glucosidases can be assigned to family 4 of glycosyl-hydrolases (see http://www.expasy.ch/cgi-bin/lists?glycosid.txt and http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html). By their composition and modular structure, proteins MalB, GlvA and AglA, can be assigned to the $EII^{Glc/Ser}$ family of PTS transporters (Lengeler *et al.*, 1994; Lanz & Erni, 1998).

Growth studies with *F. mortiferum*, *K. pneumoniae* and *E. coli* O157:H7

Growth studies were performed to determine if possession of the putative operons would permit growth of the three organisms on sucrose isomers and other α-glucosides. Both *K. pneumoniae* and *F. mortiferum* showed excellent growth on all sugars tested, including glucose, fructose, methyl α-glucoside, maltose, maltitol, sucrose and all five α-D-glucosyl-D-fructoses. *E. coli* O157:H7 grew well on glucose, fructose, maltose and

sucrose, but the pathogen was unable to grow on methyl α-glucoside, maltitol or any of the sucrose isomers.

Enzyme expression during growth of *F. mortiferum* on sucrose and its isomers

S6PH and ATP-dependent fructokinase (FK) are induced by growth of *F. mortiferum* on sucrose (Robrish *et al*., 1991; Thompson *et al*., 1992). Because sucrose and its isomers comprise the same hexose moieties, it was of interest to determine whether sucrose-grown cells of *F. mortiferum* would also metabolize the isomeric compounds. As expected, sucrose-grown cells readily fermented sucrose, but there was no detectable metabolism of the five isomers, including palatinose and turanose (Fig. 3). Furthermore, whereas sucrose-grown cells of *F. mortiferum* contained high levels of S6PH and FK activity (Table 1), after growth on the isomers, the activities of the two enzymes were not significantly greater than the constitutive levels found in glucose- or fructose-grown cells. These findings contrast markedly with the high levels of S6PH and FK that are expressed during growth of *K. pneumoniae* on the isomeric compounds (Table 2).

Phospho-α-glucosidase (MalH) is expressed during growth of *F. mortiferum* on sucrose isomers

From the results of fermentation studies and enzymic analyses (Fig. 3 and Table 1, respectively), it was evident that that dissimilation of the sucrose isomers by *F. mortiferum* was via a route that was separate from that encoded by the *scr* regulon. Because previous studies with *K. pneumoniae* (Thompson *et al*., 2001a) showed that growth on the sucrose isomers induced high-level expression of phospho-α-glucosidase (AglB), the various cell extracts of *F. mortiferum* were accordingly assayed for phospho-α-glucosidase (MalH) activity (Table 1). Extracts prepared from organisms grown previously on the sucrose isomers and other α-glucosides (e.g. maltose and methyl α-glucoside) readily hydrolysed pNP-α-G6P (Table 1). However, there was no detectable hydrolysis of the chromogenic analogue by similarly prepared extracts from organisms grown on glucose, fructose or sucrose. The results of a Western blot (Fig. 4A), performed with polyclonal antibody to MalH, confirmed expression of the phospho-α-glucosidase (molecular mass ~ 50 kDa) during growth on sucrose isomers and other α-glucosides. Significantly, the immunoreactive protein (MalH) was not detectable in either sucrose- or glucose-grown cell extracts. The data presented in Fig. 4(B) established the co-identity of the immunoreactive polypeptide and the enzymically active protein. In this experiment, samples of the various cell extracts of *F. mortiferum* were electrophoresed under non-denaturing conditions, prior to *in situ* staining for phospho-α-glucosidase activity using the fluorogenic substrate 4-MU-α-G6P. The zymogram (Fig. 4B) yielded three significant results: (i) the intensely fluorescent aglycone (4-methylumbelliferone) was generated only by those extracts that contained the immunoreactive protein (MalH); (ii) formation of a single zone of fluorescence (at the same migration distance in the gel) provided evidence for only one species of phospho-α-glucosidase in the extracts, and (iii) absence of fluorescence in lane 2 was consistent with the inability of *F. mortiferum* to express MalH during growth on sucrose.

Table 3. Cofactor requirements for the hydrolysis of chromogenic substrates pNP-α-G6P and pNP-β-G6P by purified MalH from *F. mortiferum*

Assay composition and procedures are described in Methods. NAD^+, metal ions and chromogenic substrates were present at 1 mM final concentration. Hydrolysis rates, expressed as μmol pNP formed min^{-1} (mg protein)$^{-1}$, are the means of two determinations. ND, No detectable activity.

Additions to assay	Activity with chromogenic analogue:	
	pNP-α-G6P	pNP-β-G6P
None	ND	ND
NAD^+	0·11	0·08
Mn^{2+}	1·61	1·19
$NAD^+ + Mn^{2+}$	2·91	1·56
$NAD^+ + Mg^{2+}$	0·39	0·16
$NAD^+ + Ni^{2+}$	0·51	0·54
$NAD^+ + Co^{2+}$	1·47	2·28

Purification, cofactor requirements, and substrate specificity of MalH

The data presented thus far (although suggestive), did not establish a functional role for MalH in dissimilation of the α-D-glucosyl-D-fructoses by *F. mortiferum*. Clearly, it was necessary to purify MalH, and demonstrate hydrolysis of either free or phosphorylated derivatives of the isomers by the enzyme. To this end a plasmid (pCB4.11) containing the cloned *malH* gene under its own promoter (Bouma *et al*., 1997) was transformed into *E. coli* strain PEP43. Importantly, the latter strain is deficient in all phospho-β-glucosidase activities, and an extract of these cells is unable to hydrolyse the chromogenic substrate pNP-β-G6P. Unexpectedly, after expression of MalH in *E. coli* PEP43(pCB4.11), the resultant cell extract caused the hydrolysis of both pNP-α- and pNP-β-G6P. Hydrolysis of both compounds was observed throughout purification of MalH (Fig. 4C), and the same cofactors (divalent metal ion and NAD^+) were required for cleavage of both chromogenic substrates by electrophoretically pure enzyme (Table 3). We conclude that a single protein (MalH) is responsible for the hydrolysis of the two stereomers, pNP-α-G6P and pNP-β-G6P. Studies of substrate specificity revealed that neither purified MalH, nor extracts prepared from cells of *F. mortiferum* grown previously on the isomers, were

Table 4. Hydrolysis of phosphorylated sucrose isomers by MalH from *F. mortiferum*

Assay procedures are described in Methods. Phosphorylated compounds were present at a concentration of 2 mM. Enzyme activity is expressed as μmol G6P formed min^{-1} (mg protein)$^{-1}$; values are means of two determinations. ND, No detectable hydrolysis.

Disaccharide phosphate in assay	Specific activity
Trehalulose-6′P*	0·15
Sucrose-6P	ND
Turanose-6′P*	0·68
Maltulose-6′P*	0·40
Leucrose-6′P*	0·08
Palatinose-6′P*	0·11
Maltose-6′P	0·15
Cellobiose-6′P†	ND
Gentiobiose-6′P‡	ND

* Sucrose isomers.

† Cellobiose: 4-*O*-*β*-D-glucopyranosyl-D-glucopyranose.

‡ Gentiobiose: 6-*O*-*β*-D-glucopyranosyl-D-glucopyranose.

able to hydrolyse the free (non-phosphorylated) forms of the isomeric compounds (data not shown). However, the 6′-*O*-phosphorylated derivatives of the five α-D-glucosyl-fructoses were hydrolysed by MalH (Table 4), albeit at a rate considerably slower than that determined for the α-linked chromogenic substrate, pNP-α-G6P. Significantly, sucrose 6-phosphate itself was not a substrate for MalH, and the enzyme also failed to hydrolyse *β*-*O*-linked phosphorylated disaccharides such as cellobiose 6′-phosphate and gentiobiose 6′-phosphate (Table 4).

DISCUSSION

Here we report the metabolism of sucrose isomers by *F. mortiferum*, and provide insight into the enzymic and genetic basis for growth on these isomers. Presently, *F. mortiferum* and *K. pneumoniae* are the only organisms known to ferment the five α-D-glucosyl-D-fructoses. Earlier we showed that MalH from *F. mortiferum* hydrolysed maltose 6′-phosphate (Thompson *et al.*, 1995), and now we provide evidence for the cleavage of the phosphorylated isomers of sucrose by this enzyme. The phospho-α-glucosidase gene (*malH*) is adjacent to the gene *malB*, whose now completed sequence predicts a polypeptide that in size, domain structure, and conserved motifs (GITE and CATRLR) is characteristic of an EII(CB) transporter of the PEP:PTS. Genes *malB* and *malH* are homologous to *aglA* and *aglB*, respectively, of *K. pneumoniae*. We suggest that the polypeptides encoded by these genetic elements are required for growth of *F. mortiferum* and *K. pneumoniae* on the isomers of sucrose and related α-glucosides, including maltose. A common feature of these genetic units is the absence of a gene encoding a third (and usually sugar-specific) protein (EIIA) that is required for operation of all PTS systems. Interestingly, both sucrose:PTS and trehalose:PTS operons in *Bacillus subtilis* also lack the expected EIIA genes and, for these systems, it is believed that $EIIA^{G1c}$ can serve as substitute (Sutrina *et al.*, 1990; Dahl, 1997). A similar cross-complementation may also occur between the $EIIA^{G1c}$ and EII(CB) proteins of *F. mortiferum* and *K. pneumoniae* to yield a functional α-glucoside:PTS in these species.

Proteins encoded by the *scr* operons of *K. pneumoniae* and *F. mortiferum* are expressed during growth of both organisms on sucrose (Sprenger & Lengeler, 1988; Thompson *et al.*, 1992). Hydrolysis of sucrose 6-phosphate by S6PH yields G6P and fructose, and for *K. pneumoniae*, fructose is believed to be the inducer of the *scr* operon (Jahreis & Lengeler, 1993). Hydrolysis of the phosphorylated isomers by AglB of *K. pneumoniae* also yields G6P and fructose, and formation of the latter ketohexose is consistent with the high levels of S6PH and FK present in cells grown on the isomers (Table 2). Surprisingly, similar studies with *F. mortiferum* showed that, for this organism, growth on the isomeric compounds did not induce significant expression of either S6PH or FK (Table 1). These findings explain why these cells were unable to metabolize sucrose, and additionally, the data point to sucrose 6-phosphate (rather than fructose) as the likely inducer of the *scr* operon in *F. mortiferum*.

Substrate specificity and hydrolysis of chromogenic substrates by MalH

MalH is an oligomeric protein comprising four identical subunits (molecular mass ~ 50 kDa) that, by sequence-based alignment, is assigned to family 4 of the glycosylhydrolase superfamily. As reported for other members of this unusual family, MalH is inherently unstable and Mn^{2+} and NAD^{+} are prerequisite cofactors for activity (Nagao *et al.*, 1988; Thompson *et al.*, 1998, 1999; Raasch *et al.*, 2000). Whether the nucleotide and metal ion fulfil catalytic or structural functions has not been ascertained for any member of family 4. Phosphorylation at O-6 of the glucosyl moiety of the isomers is necessary for substrate cleavage, and MalH is unable to hydrolyse the corresponding non-phosphorylated compounds. MalH is also exacting with respect to the α-*O* linkage of its PTS-derived substrates (see below) and, because there is no detectable hydrolysis of *β*-*O*-linked stereomers such as cellobiose 6′-phosphate and gentiobiose 6′-phosphate (Table 4), the enzyme may reasonably be classified as a phospho-α-glucosidase. In this context, it is not clear why MalH should hydrolyse both pNP-α-G6P and pNP-*β*-G6P with comparable efficiency (Table 3). Co-purification of MalH with a phospho-*β*-glucosidase resident in the host (*E. coli* PEP43) can be discounted because of gene inactivation or crypticity, and analysis of the final preparation by SDS-PAGE provided evidence for only a single polypeptide. That the same cofactors should also be required for catalysis is further evidence that the same enzyme hydrolyses both α- and *β*-forms of the chromogenic compound(s).

Unlike the phosphorylated sucrose isomers (where G6P is linked to a fructose moiety), the essential G6P moiety of the chromogenic substrates is attached to *p*-nitrophenol. Perhaps the aromatic aglycone exerts an effect (electron-withdrawing?) upon the *O*-linkage such that α/β conformation is no longer a determinant of substrate specificity. It is of comparative interest to note that cellobiose-6-phosphate hydrolase (CelF) from *E. coli* is also a member of glycosylhydrolase family 4 (Thompson *et al.*, 1999). In contrast to MalH, this NAD^+- and metal-dependent phospho-β-glucosidase hydrolyses only pNP-β-G6P.

Molecular basis for substrate recognition by MalH

Sucrose 6-phosphate and its phosphorylated isomers are not commercially available, but all of these derivatives were recently prepared in our laboratory (Thompson *et al.*, 2001a). Studies of substrate specificity showed that whereas MalH hydrolysed all of the phosphorylated isomers, the enzyme failed to hydrolyse sucrose 6-phosphate itself. Insight into the molecular basis for this remarkable discrimination among potential substrates was gained by molecular dynamics simulations, which revealed the probable solution-state geometries of the various disaccharide phosphates (Thompson *et al.*, 2001b). Molecular dynamics simulations and determination of solvent-accessible surfaces indicate pronounced conformational differences between sucrose 6-phosphate and its five isomeric 6′-phosphates. By virtue of an interresidue water bridge between Glc-2-O$\cdots H_2O\cdots$O-1-Fru, both sucrose and sucrose 6-phosphate assume a compact, globular shape in solution (Immel & Lichtenthaler, 1995; Thompson *et al.*, 2001b). This water bridge is not present in the 6′-phosphoglucosyl-fructoses and, in consequence, the phosphorylated isomers adopt a more linear (extended) molecular geometry. The specificity of MalH for the isomeric phosphates presumably reflects recognition by the enzyme's binding domain of both the shape and the molecular lipophilicity potential of the contact surfaces of these particular molecules (Thompson *et al.*, 2001b).

Conclusions

Both *F. mortiferum* and *K. pneumoniae* readily metabolize the five isomers of sucrose. In contrast, *E. coli* O157:H7 (which grows well on sucrose) failed to grow on any of the isomeric compounds. These results were surprising, because this enterohaemorrhagic strain has three genes (*yidP*, *glvA* and *glvG*) whose organization and deduced amino acid sequences are virtually identical to those of *aglR*, *aglA* and *aglB*, respectively, in *K. pneumoniae*. Although contrary to expectation, the results obtained for *E. coli* O157:H7 were nevertheless important. First, the data established for *E. coli* O157:H7 (as for the other species), that the sucrose-PTS/S6PH pathway is neither induced by, nor does it provide a route for dissimilation of, sucrose isomers. Secondly, the data indicate that possession of genes encoding α-glucoside-specific EII(CB) and phospho-α-glucosidase (while necessary), may not be entirely sufficient for dissimilation of α-D-glucosyl-D-fructoses by micro-organisms.

ACKNOWLEDGEMENTS

We thank Carolyn L. Bouma for construction and provision of pCB4.11. We would like to thank Jack London and Edith C. Wolff for their review and constructive criticisms of this article. We express appreciation to Professor Frieder W. Lichtenthaler for his encouragement and interest in our investigation.

REFERENCES

Bouma, C. L., Reizer, J., Reizer, A., Robrish, S. A. & Thompson, J. (1997). 6-Phospho-α-D-glucosidase from *Fusobacterium mortiferum*: cloning, expression, and assignment to family 4 of the glycosylhydrolases. *J Bacteriol* **179**, 4129–4137.

Chen, Y.-Y. M. & LeBlanc, D. J. (1992). Genetic analysis of *scrA* and *scrB* from *Streptococcus sobrinus* 6715. *Infect Immun* **60**, 3739–3746.

Dahl, M. K. (1997). Enzyme II^{Glc} contributes to trehalose metabolism in *Bacillus subtilis*. *FEMS Microbiol Lett* **148**, 233–238.

Fouet, A., Arnaud, M., Klier, A. & Rapoport, G. (1987). *Bacillus subtilis* sucrose-specific enzyme II of the phosphotransferase system: expression in *Escherichia coli* and homology to enzymes II from enteric bacteria. *Proc Natl Acad Sci U S A* **84**, 8773–8777.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* **280**, 309–316.

Immel, S. & Lichtenthaler, F. W. (1995). Molecular modeling of saccharides. 7. The conformation of sucrose in water: a molecular dynamics approach. *Liebigs Ann Chem* 1925–1937.

Jahreis, K. & Lengeler, J. W. (1993). Molecular analysis of two ScrR repressors and of a ScrR-FruR hybrid repressor for sucrose and D-fructose specific regulons from enteric bacteria. *Mol Microbiol* **9**, 195–209.

Lanz, R. & Erni, B. (1998). The glucose transporter of the *Escherichia coli* phosphotransferase system. Mutant analysis of the invariant arginines, histidines, and domain linker. *J Biol Chem* **273**, 12239–12243.

Lengeler, J. W., Jahreis, K. & Wehmeier, U. F. (1994). Enzymes II of the phospho*enol*pyruvate-dependent phosphotransferase systems: their structure and function in carbohydrate transport. *Biochim Biophys Acta* **1188**, 1–28.

Lichtenthaler, F. W. & Rönninger, S. (1990). α-D-Glucopyranosyl-D-fructoses: distribution of furanoid and pyranoid tautomers in water, dimethyl sulphoxide, and pyridine. Studies on ketoses. Part 4. *J Chem Soc Perkin Trans* **2**, 1489–1497.

Lichtenthaler, F. W., Immel, S. & Kreis, U. (1991). Evolution of the structural representation of sucrose. *Starch/Stärke* **43**, 121–132.

Loesche, W. J. (1986). Role of *Streptococcus mutans* in human dental decay. *Microbiol Rev* **50**, 353–380.

Meadow, N. D., Fox, D. K. & Roseman, S. (1990). The bacterial phosphoenolpyruvate:glycose phosphotransferase system. *Annu Rev Biochem* **59**, 497–542.

Minami, T., Fujiwara, T., Ooshima, T., Nakajima, Y. & Hamada, S. (1990). Interaction of structural isomers of sucrose in the reaction between sucrose and glucosyltransferases from mutans streptococci. *Oral Microbiol Immunol* **5**, 189–194.

Nagao, Y., Nakada, T., Imoto, M., Shimamoto, T., Sakai, S., Tsuda, M. & Tsuchiya, T. (1988). Purification and analysis of the structure

of α-galactosidase from *Escherichia coli*. *Biochem Biophys Res Commun* **151**, 236–241.

Ooshima, T., Izumitani, A., Sobue, S., Okahashi, N. & Hamada, S. (1983). Non-cariogenicity of the disaccharide palatinose in experimental dental caries of rats. *Infect Immun* **39**, 43–49.

Ooshima, T., Izumitani, A., Minami, T., Fujiwara, T., Nakajima, Y. & Hamada, S. (1991). Trehalulose does not induce dental caries in rats infected with mutans streptococci. *Caries Res* **25**, 277–282.

Peltroche-Llacsahuanga, H., Hauk, C. J., Kock, R., Lampert, F., Lütticken, R. & Haase, G. (2001). Assessment of acid production by various human oral micro-organisms when palatinose or leucrose is utilized. *J Dent Res* **80**, 378–384.

Perna, N. T., Plunkett, G., III, Burland, V. & 25 other authors (2001). Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**, 529–533.

Postma, P. W., Lengeler, J. W. & Jacobson, G. R. (1993). Phospho*enol*pyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev* **57**, 543–594.

Raasch, C., Streit, W., Schanzer, J., Bibel, M., Gosslar, U. & Liebl, W. (2000). *Thermotoga maritima* AglA, an extremely thermostable NAD^+-, Mn^{2+}-, and thiol-dependent α-glucosidase. *Extremophiles* **4**, 189–200.

Rauch, P. J. G. & deVos, W. M. (1992). Transcriptional regulation of the Tn*5276*-located *Lactococcus lactis* sucrose operon and characterization of the *sacA* gene encoding sucrose-6-phosphate hydrolase. *Gene* **121**, 55–61.

Reid, S. J., Rafudeen, M. S. & Leat, N. G. (1999). The genes controlling sucrose utilization in *Clostridium beijerinckii* NCIMB 8052 constitute an operon. *Microbiology* **145**, 1461–1472.

Robrish, S. A. & Thompson, J. (1990). Regulation of fructose metabolism and polymer synthesis by *Fusobacterium nucleatum* ATCC 10953. *J Bacteriol* **172**, 5714–5723.

Robrish, S. A., Oliver, C. & Thompson, J. (1987). Amino acid-dependent transport of sugars by *Fusobacterium nucleatum* ATCC 10953. *J Bacteriol* **169**, 3891–3897.

Robrish, S. A., Oliver, C. & Thompson, J. (1991). Sugar metabolism by fusobacteria: regulation of transport, phosphorylation, and polymer formation by *Fusobacterium mortiferum* ATCC 25557. *Infect Immun* **59**, 4547–4554.

Schmid, K., Ebner, R., Altenbuchner, J., Schmitt, R. & Lengeler, J. (1988). Plasmid-mediated sucrose metabolism in *Escherichia coli* K12: mapping of the *scr* genes of pUR400. *Mol Microbiol* **2**, 1–8.

Slee, A. M. & Tanzer, J. M. (1979). Phosphoenolpyruvate-dependent sucrose phosphotransferase activity in *Streptococcus mutans* NCTC 10449. *Infect Immun* **24**, 821–828.

Sprenger, G. A. & Lengeler, J. W. (1988). Analysis of sucrose catabolism in *Klebsiella pneumoniae* and in Scr^+ derivatives of *Escherichia coli* K12. *J Gen Microbiol* **134**, 1635–1644.

St Martin, E. J. & Wittenberger, C. L. (1979). Characterization of a phosphoenolpyruvate-dependent sucrose phosphotransferase system in *Streptococcus mutans*. *Infect Immun* **24**, 865–868.

Sutrina, S. L., Reddy, P., Saier, M. H., Jr & Reizer, J. (1990). The glucose permease of *Bacillus subtilis* is a single polypeptide chain that functions to energize the sucrose permease. *J Biol Chem* **265**, 18581–18589.

Tangney, M., Rousse, C., Yazdanian, M. & Mitchell, W. J. (1998). Sucrose transport and metabolism in *Clostridium beijerinkii* NCIMB 8052. *J Appl Microbiol* **84**, 914–919.

Thompson, J. & Chassy, B. M. (1981). Uptake and metabolism of sucrose by *Streptococcus lactis*. *J Bacteriol* **147**, 543–551.

Thompson, J., Nguyen, N. Y., Sackett, D. L. & Donkersloot, J. A. (1991). Transposon-encoded sucrose metabolism in *Lactococcus lactis*. Purification of sucrose-6-phosphate hydrolase and genetic linkage to N^5-(L-1-carboxyethyl)-L-ornithine synthase in strain K1. *J Biol Chem* **266**, 14573–14579.

Thompson, J., Nguyen, N. Y. & Robrish, S. A. (1992). Sucrose fermentation by *Fusobacterium mortiferum* ATCC 25557: transport, catabolism, and products. *J Bacteriol* **174**, 3227–3235.

Thompson, J., Gentry-Weeks, C. R., Nguyen, N. Y., Folk, J. E. & Robrish, S. A. (1995). Purification from *Fusobacterium mortiferum* of a 6-phosphoryl-O-α-D-glucopyranosyl:6-phosphoglucohydrolase that hydrolyzes maltose 6-phosphate and related phospho-α-D-glucosides. *J Bacteriol* **177**, 2505–2512.

Thompson, J., Pikis, A., Ruvinov, S. B., Henrissat, B., Yamamoto, H. & Sekiguchi, J. (1998). The gene *glvA* of *Bacillus subtilis* 168 encodes a metal-requiring, NAD(H)-dependent 6-phospho-α-glucosidase. Assignment to family 4 of the glycosylhydrolase superfamily. *J Biol Chem* **273**, 27347–27356.

Thompson, J., Ruvinov, S. B., Freedberg, D. I. & Hall, B. G. (1999). Cellobiose-6-phosphate hydrolase (CelF) of *Escherichia coli*: characterization and assignment to the unusual family 4 of glycosylhydrolases. *J Bacteriol* **181**, 7339–7345.

Thompson, J., Robrish, S. A., Pikis, A., Brust, A. & Lichtenthaler, F. W. (2001a). Phosphorylation and metabolism of sucrose and its five linkage-isomeric α-D-glucosyl-D-fructoses by *Klebsiella pneumoniae*. *Carbohydr Res* **331**, 149–161.

Thompson, J., Robrish, S. A., Immel, S., Lichtenthaler, F. W., Hall, B. G. & Pikis, A. (2001b). Metabolism of sucrose and its five linkage-isomeric α-D-glucosyl-D-fructoses by *Klebsiella pneumoniae*: participation and properties of sucrose-6-phosphate hydrolase and phospho-α-glucosidase. *J Biol Chem* **276**, 37415–37425.

Titgemeyer, F., Jahreis, K., Ebner, R. & Lengeler, J. W. (1996). Molecular analysis of the *scrA* and *scrB* genes from *Klebsiella pneumoniae* and plasmid pUR400, which encode the sucrose transport protein enzyme II^{Scr} of the phosphotransferase system and a sucrose-6-phosphate invertase. *Mol Gen Genet* **250**, 197–206.

Van Houte, J. (1994). Role of microorganisms in caries etiology. *J Dent Res* **73**, 672–681.

Ziesenitz, S. C., Siebert, G. & Imfeld, T. (1989). Cariological assessment of leucrose [D-glucopyranosyl-α(1–5)-D-fructopyranose] as a sugar substitute. *Caries Res* **23**, 351–357.

Received 3 September 2001; revised 5 November 2001; accepted 8 November 2001.