



- (51) International Patent Classification:
C07K 14/395 (2006.01)
- (21) International Application Number:
PCT/EP2014/061640
- (22) International Filing Date:
4 June 2014 (04.06.2014)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
13170648.3 5 June 2013 (05.06.2013) EP
- (71) Applicant: **DSM IP ASSETS B.V.** [NL/NL]; Het Overloon 1, NL-6411 TE Heerlen (NL).
- (72) Inventors: **KLAASSEN, Paul**; P.O. Box 4, NL-6100 AA Echt (NL). **DE WAAL, Paulus Petrus**; P.O. Box 4, NL-6100 AA Echt (NL). **DE JONG, René Marcel**; P.O. Box 4, NL-6100 AA Echt (NL).
- (74) Agent: **KLEIBORN, Paul Erik**; DSM Intellectual Property, P.O. Box 4, NL-6100 AA Echt (NL).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,

BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: POLYPEPTIDES WITH PERMEASE ACTIVITY

(57) Abstract: The invention relates to a polypeptide having one or more substitution at a position corresponding to position 339 or 376 of SEQ ID NO: 59, wherein the polypeptide is a member of the Major Facilitator Superfamily (MFS). In an embodiment, the substitution is at position corresponding to 376 and wherein the aminoacid at that position is replaced by an aminoacid that has a van der Waals volume of 80 to 138 Å³ and a side chain hydrophobicity of 10 to 100 Δtr.



WO 2014/195381 A2

POLYPEPTIDES WITH PERMEASE ACTIVITY

Field of the invention

The invention is directed to novel polypeptides and to recombinant
5 organisms expressing the polypeptides. In an embodiment, the present invention relates
to novel permease polypeptides with altered sugar specificity and/or sugar transport
activity.

Background of the invention

10 The plasma membrane of yeast cells and other eukaryotes is a complex bio-
membrane, consisting of two layers of phospholipids, with a plethora of proteins
embedded in it. Many molecules may cross the plasma membrane by diffusion and
osmosis or with the aid of specific transport systems.

15 Transport systems allow the uptake of nutrients and ions, export of products of
metabolism and undesired or harmful substances. Different mechanisms exist. Primary
active transporters drive solute accumulation or extrusion by using for instance ATP
hydrolysis. Secondary carriers, belonging to the Major Facilitator Superfamily (MFS)
transporters, facilitate the transport of one or more molecular species across the
20 membrane in response to chemi-osmotic gradients. In the yeast *Saccharomyces
cerevisiae*, 186 MFS proteins have been identified (Nelissen, 1997) in strain S288c.

An example of such a carrier is the Hxt1 protein, involved in hexose transport in
Saccharomyces cerevisiae.

25 Permeases are membrane transport proteins, a class of multipass
transmembrane proteins that facilitate the diffusion of a specific molecule, herein
specifically one or more sugar, in or out of the cell by passive transport. In contrast,
active transporters couple molecule transmembrane transport with an energy source
such as ATP or a favorable ion gradient.

30 The terms permease, facilitator, transporter or transport protein or related terms
are all describing proteins with multiple membrane spanning domains that exhibit a
function in transporting molecules across a membrane. This transport can be brought
about by different mechanisms: uniport (transport of one molecule), symport
(simultaneous co-transport of two different molecules in the same direction), antiport
(simultaneous transport of two molecules in opposite directions) and facilitated diffusion.

The family of sugar transporters in yeast consists of 30-40 members (34 members in strain S288c (Nelissen, 1997)). The sugar transporters can be divided in five clusters: hexose permeases (*HXT*-genes, *GAL2*), disaccharide permeases, myo-inositol permeases, sugar receptors and a final cluster of transporters of which the substrate is unknown.

Lignocellulosic biomass, an attractive alternative feedstock for the production of liquid transportation fuels, consists of several different sugars. The hexose fraction of lignocellulose, mainly glucose, can in principle be readily fermented by non-recombinant versions of the yeast *S. cerevisiae*. However, this organism is not able to metabolize the pentose sugars, such as xylose and arabinose, into ethanol.

Methods of creating microorganisms that are able to metabolize pentose sugars are known in the art. For instance, in WO/2009/109630 the construction of expression cassettes and the transformation of *S. cerevisiae* cells into pentose fermenting cells by expressing xylose isomerase are illustrated.

Native pentose-utilizing organisms exist but are lacking well-developed genetic tools and/or low product tolerances, which limit their suitability as hosts for lignocellulosic conversion processes.

As a consequence, efforts have focused on the introduction of pentose conversion pathways in the yeast *S. cerevisiae*, which is still the organism of choice in the ethanol industry, in order to enable pentose fermentation.

Despite the vast amount of progress achieved in the past years, the transport of pentose sugars is still considered to be (one of) the rate-limiting step in pentose metabolism.

Pentose transport in *S. cerevisiae* is mediated by the different members of the hexose transporter (Hxt) family. Hxt4, Hxt5, Hxt7 and Gal2 have been described as the main xylose transporters in *S. cerevisiae* (Hamacher et al, 2002), and Gal2 is also known to mediate arabinose transport (Becker et al, 2003). However, the affinity for the respective pentose sugars is approximately 10 to 100 times lower than for the respective hexose sugars. The lack of a dedicated xylose or arabinose transporter in recombinant yeast cells thus limits the capacity for co-utilization of hexoses and pentoses in sugar mixtures, and prohibits a high pentose catabolic flux. As a consequence, conversion of biomass sugars may be considered bi-phasic: in the first phase, a relatively fast conversion of hexoses (glucose) takes place, while in the second phase, which starts

when the hexoses have been exhausted from the medium, pentose fermentation commences, but at a far lower rate as compared to the rate of hexose conversion.

It is therefore a long-felt desire to express pentose-specific sugar transporters, i.e. no glucose interference (pentose specificity) and high affinity to pentose, in an otherwise unchanged transporter landscape, in order to maintain the ability to convert
5 hexoses at approximately the same level.

One way of solving this problem is to screen for heterologous sugar (pentose) transporters which are pentose specific and have a (moderately) high affinity for pentose. However, such efforts have been with limited success so far. Only a few have
10 been shown to be able to facilitate pentose transport upon expression in *S. cerevisiae*, but all favour glucose above xylose (Young *et al*, 2011, and references therein), as is the case with the *S. cerevisiae* Hxt-proteins, as indicated above.

Another approach is to re-engineer hexose transporters to pentose transporters. For instance, the works by Kasahara *et al* (2000; 2009; 2010) indicate which residues in
15 several sugar transporters play a key role in the determination of the substrate affinity to the natural substrate.

Mutant hexose transporters that are able to transport pentose sugars more efficiently are known in the art. For instance, in WO/2012/049173, the isolation of mutant hexose transporter genes from cultures of pentose-fermenting *S. cerevisiae* cells is
20 described.

In *Saccharomyces cerevisiae*, the permease GAL2 transports galactose across the cell membrane. It is also known as a transporter of glucose across the membrane.

Summary of the invention

25 An object of the invention is to provide novel permease polypeptides with altered, in particular improved, sugar specificity. Another object of the invention is to provide recombinant strains expressing the permease polypeptide that have improved uptake of the molecule that the permease transports across the cell membrane. Another object is to provide a permease polypeptide that has a improved capacity for transport of C5
30 sugars, in particular xylose compared to a parent polypeptide. Another object is to provide a permease polypeptide that has reduced transport activity for C6 sugar, in particular glucose, compared to a parent polypeptide. Another object is to provide a method to identify mutations in other related permease polypeptides that have a

beneficial effect on the improved capacity for transport of xylose or reduced transport activity for glucose.

One or more of these objects are attained according to the invention. According to the present invention, there is provided a polypeptide having one or more substitution at a position corresponding to position 339 or 376 of SEQ ID NO: 59, wherein the polypeptide is a member of the Major Facilitator Superfamily (MFS). In an embodiment, the substitution is at position corresponding to 376 and wherein the aminoacid at that position is replaced by an aminoacid that has a van der Waals volume of 80 to 138 Å³ and a side chain hydrophobicity of 10 to 100 Δt_R. In an embodiment the amino acid has at position 376 have a van der Waals volume of 85 to 138 Å³ and a side chain hydrophobicity of 10 to 100 Δt_R.

The values for van der Waals volume (Å³) for aminoacids are herein used as described in: http://www.proteinsandproteomics.org/content/free/tables_1/table08.pdf. The corresponding literature is N. J. Darby, Thomas E. Creighton, Protein Structure (1993) Oxford University Press. The values for side chain hydrophobicity (Δt_R) of aminoacids are herein used as described in <http://onlinelibrary.wiley.com/doi/10.1002/psc.310010507/pdf>. The reference corresponding to this is Monera, O.D. et al, Journal of Peptide science Vol. 1, 319-329 (1995).

A polypeptide according to the invention having one or more of these mutations has an advantageous sugar consumption and/or fermentation product production. This will be described in more detail below and will be illustrated by examples 1-5 below.

Brief description of the drawings

FIG. 1 shows results of aerobic shake flask cultures hexose transporter mutants on Verduyn-urea + 15 g l⁻¹ glucose + 20 g l⁻¹ xylose. (A) Optical density measurements at 600 nm wavelength, (B) glucose concentrations (g l⁻¹), (C) xylose concentrations during the culture period.

FIG. 2 shows results of aerobic shake flask cultures hexose transporter mutants on Verduyn-urea + 20 g l⁻¹ xylose. (A) Optical density measurements at 600 nm wavelength, (B) xylose concentrations (g l⁻¹).

FIG. 3 shows the strain construction scheme for strain YD01227

FIG. 4 shows results of aerobic shake flask cultures of quadruple hexose kinase mutant (col 2) and reference strain RN1014 on Verduyn-urea + 100 g l⁻¹ glucose + 20 g l⁻¹ xylose. A) Optical density measurements at 600 nm wavelength, (B) glucose concentrations (g l⁻¹), (C) xylose concentrations during the culture period.

FIG. 5 shows results of micro-well plate cultures on glucose - xylose mixtures. Growth characteristics of reference strain RN1014, and quadruple hexokinase mutants YD01227 and colony 2 (Col 2) with equal genotype on Verduyn-urea supplemented with mixtures of glucose and xylose in the following concentrations: (A) 60 + 20, (B) 100 + 20, (C) 15 + 5, and (D) 25 g l⁻¹ + 5 g l⁻¹, respectively. Every 15 minutes, OD600 was measured automatically by Bioscreen C apparatus. Data points are the average of measurements in triplicate.

FIG. 6 shows a plasmid map pRN993.

FIG. 7 shows the relationships between residual xylose and glucose concentrations (A), or residual xylose concentrations and growth (OD600; B) in the Glucose Transport Activity Counter Screen (GTAC screen), see examples. Residual xylose concentrations (g l⁻¹) were plotted against the residual (A) glucose concentration (g l⁻¹) and (B) OD600 at 96 hours. (Inset B) linear regression analysis of shows a good correlation of OD600 and xylose. Each data point represents the average of sugar concentrations of 1 to 3 replicates, or of multiple measurements of RN1001 control strain or medium samples per part of the screen.

FIG. 8 shows growth and xylose consumption of Gal2-N376-mutant variants in the GTAC screen. (A) OD600 measurements at time points 24, 72 and 96 hours, and (B) residual xylose concentrations (g l⁻¹) measured for the Gal2-N376-mutant variants, RN1001 control strain and empty samples (YD10 medium-only) during the GTAC Screen. Each column represents the average OD600 or residual xylose concentration of 3 transformants; error bars represent the standard error of the mean. Asterisks indicate amino acids with large hydrophobic side chains with a clear effect.

FIG. 9 shows growth profiles of TOP15 Xylose Transport Activity (XTA) Screening (see examples). OD600 values of the three sampling points for the TOP15 strains/variants from (A) Part A and (B) Part B of the XTA Screen. Each column

represents the average OD600 of 3 transformants; error bars represent the standard error of the mean.

FIG. 10 shows an alignment of permease protein sequences.

FIG. 11 shows plot of amino acid van der Waals volume ($\text{\AA}^3$) (Y-axis) against amino acid side chain hydrophobicity (Δt_R) (X-axis). Advantageous amino acids (reduced glucose transport activity) for position corresponding to 376 in SEQ ID NO: 59 are within the area defined by the short striped lines and for position 339 by the long strips lines. Values for van der Waals volume used herein are described in: http://www.proteinsandproteomics.org/content/free/tables_1/table08.pdf. The

corresponding literature is N. J. Darby, Thomas E. Creighton, Protein Structure (1993) Oxford University Press. The values for side chain hydrophobicity (Δt_R) of amino acids are herein used as described in <http://onlinelibrary.wiley.com/doi/10.1002/psc.310010507/pdf>. The reference corresponding to this is Monera, O.D. et al, Journal of Peptide science Vol. 1 319-329 (1995).

Brief description of the sequence listing

SEQ ID NO: 1: primer 5034-kanf

SEQ ID NO: 2: primer 5035-kanr

SEQ ID NO: 3: primer 5116-lf2

SEQ ID NO: 4: primer 5118-lr2

SEQ ID NO: 5: primer 5115-lf1

SEQ ID NO: 6: primer 5117-lr1

SEQ ID NO: 7: pRN201; TOPO-BLUNT-loxP-kanMX-loxP

SEQ ID NO: 8: pRN251; TOPO-BLUNT-loxP-hphMX-loxP

SEQ ID NO: 9: pRN365; TOPO-BLUNT-loxP-natMX-loxP

SEQ ID NO: 10: primer 115-natf

SEQ ID NO: 11: primer 116-natr

SEQ ID NO: 12: pRN447; TOPO-BLUNT-loxP-zeoMX-loxP

SEQ ID NO: 13: primer 28-H3f

SEQ ID NO: 14: primer 29-H3r

SEQ ID NO: 15: pRN247 (TOPO- BLUNT-HIS3:loxPkanMXloxP)

SEQ ID NO: 16: primer 201-Hx2uf

- SEQ ID NO: 17: primer 202-Hx2ur
SEQ ID NO: 18: primer 203-Hx2df
SEQ ID NO: 19: primer 204-Hx2dr
SEQ ID NO: 20: primer 205-Hx3uf
5 SEQ ID NO: 21: primer 206-Hx3ur
SEQ ID NO: 22: primer 210-Hx4df
SEQ ID NO: 23: primer 211-Hx4dr
SEQ ID NO: 24: primer 212-Hx5uf
SEQ ID NO: 25: primer 213-Hx5ur
10 SEQ ID NO: 26: primer 229-Hx7df
SEQ ID NO: 27: primer 230-Hx7dr
SEQ ID NO: 28: primer 243-Gal2ufn
SEQ ID NO: 29: primer 244-Gal2urn
SEQ ID NO: 30: primer 233-Ga2df
15 SEQ ID NO: 31: primer 234-Ga2dr
SEQ ID NO: 32: pRN485; TOPO-BLUNT-GAL2:loxPzeoMXloxP
SEQ ID NO: 33: pRN566; TOPO-BLUNT-HXT367:loxP-hphMX-loxP
SEQ ID NO: 34: pRN569; TOPO-BLUNT-HXT514:loxP-natMX-loxP
SEQ ID NO: 35: pRN635; TOPO-BLUNT-HXT2:loxP-kanMX-loxP
20 SEQ ID NO: 36: primer 281-Hx3inr2
SEQ ID NO: 37: primer 323-Hx7inr1
SEQ ID NO: 38: primer Hx4inr2
SEQ ID NO: 39: primer Hx5inf
SEQ ID NO: 40: primer 324-Ga2inf1
25 SEQ ID NO: 41: primer 325-Ga2inr1 SEQ ID NO 42: primer 289-Hx2inf
SEQ ID NO: 43: primer 290-Hx2inr
SEQ ID NO: 44: primer 838-Glk1-psuc227f
SEQ ID NO: 45: primer 834- Hxk2-psuc227f
SEQ ID NO: 46: primer 645-pSUC227r
30 SEQ ID NO: 47: primer 839-Glk1-psuc225r
SEQ ID NO: 48: primer 835-Hxk2-psuc225r
SEQ ID NO: 49: primer 646-pSUC225f
SEQ ID NO: 50: primer 846-Hxk1_loxP_f
SEQ ID NO: 51: primer 847-Hxk1_loxP_r

SEQ ID NO: 52: primer 848-Gal1_loxP_f

SEQ ID NO: 53: primer 849-Gal1_loxP_r

SEQ ID NO: 54: pRN774; TOPO-BLUNT-loxP-hphMX-loxP (loxP sites in opposite orientation)

5 SEQ ID NO: 55: pRN775; TOPO-BLUNT-loxP-natMX-loxP (loxP sites in opposite orientation)

SEQ ID NO: 56: WT-GAL2 DNA sequence SEQ ID NO: 57 pRN993; XbaI site (TCTAGA) and BssHII site (GCGCGC).

10 SEQ ID NO: 58: pDB1250; WT-GAL2 expression vector for screening; XbaI site (TCTAGA) and BssHII site (GCGCGC).

SEQ ID NO: 59: WT Gal2p amino acid sequence

SEQ ID NO: 60: pRN187 (pSH65-derived CRE recombinase expression vector)

SEQ ID NO: 61: pRN486 (TOPO-BLUNT-his3::loxP-natMX-loxP)

15 **Detailed description of the invention**

Throughout the present specification and the accompanying claims, the words "comprise" and "include" and variations such as "comprises", "comprising", "includes" and "including" are to be interpreted inclusively. That is, these words are intended to convey the possible inclusion of other elements or integers not specifically recited, where
20 the context allows. The articles "a" and "an" are used herein to refer to one or to more than one (i.e. to one or at least one) of the grammatical object of the article. By way of example, "an element" may mean one element or more than one element.

The invention relates to a method of identifying amino acid positions in permease polypeptides, preferably hexose permease polypeptides, more preferably hexose
25 permease polypeptides from yeast and fungi, even more preferably in *Saccharomyces cerevisiae* Hxt or Gal2 permease polypeptides, which can be mutated to alter the sugar specificity of the permease.

The invention relates to a polypeptide having one or more substitution at a position corresponding to position 339 or 376 of SEQ ID NO: 59, wherein the polypeptide
30 is a member of the Major Facilitator Superfamily (MFS). In an embodiment, the substitution is at position corresponding to 376 and wherein the amino acid at that position is replaced by an amino acid that has a van der Waals volume of 80 to 138 Å³ and a side chain hydrophobicity of 10 to 100 Δt_R (T, C, V, M, L, I, F)

In an embodiment, the substitution is at position corresponding to 376 and

wherein the aminoacid at that position is replaced by an aminoacid that has a van der Waals volume of 80 to 138 Å³ and a side chain hydrophobicity of 40 to 100 Δt_R. (C, V, M, L, I, F).

In an embodiment, the substitution is at position corresponding to 376 and
 5 wherein the aminoacid at that position is replaced by an aminoacid that has a van der Waals volume of 90 to 138 Å³ and a side chain hydrophobicity of 10 to 100 Δt_R (T, V, M, L, I, F)

In an embodiment, the substitution is at position corresponding to 376 and
 wherein the aminoacid at that position is replaced by an aminoacid that has a van der
 10 Waals volume of 100 to 138 Å³ and a side chain hydrophobicity of 60 to 100 Δt_R (V, M, L, I, F).

In an embodiment, the substitution is at position corresponding to 376 and
 wherein the aminoacid at that position is replaced by an aminoacid that has a van der
 Waals volume of 100 to 130 Å³ and a side chain hydrophobicity of 60 to 100 Δt_R (V, M, L,
 15 I).

In an embodiment, the substitution is at position corresponding to 376 and
 wherein the aminoacid at that position is replaced by an aminoacid that has a van der
 Waals volume of 120 to 130 Å³ and a side chain hydrophobicity of 60 to 100 Δt_R (M, L, I).

In an embodiment, the substitution is at position corresponding to 376 and
 20 wherein the aminoacid at that position is replaced by an aminoacid that has a van der Waals volume of 120 to 130 Å³ and a side chain hydrophobicity of 80 to 100 Δt_R (L,I).

In an embodiment, the substitution is at position corresponding to 376 and
 wherein the aminoacid at that position is replaced by an aminoacid that has a van der
 Waals volume of 100 to 130 Å³ and a side chain hydrophobicity of 60 to 80 Δt_R (V, M).

25 In an embodiment, the substitution is at position corresponding to 376 and
 wherein the aminoacid at that position is replaced by an aminoacid that has a van der Waals volume of 100 to 130 Å³ and a side chain hydrophobicity of 60 to 98 Δt_R (V, M, L).

In an embodiment, the substitution is N376T, N376C, N376V, N376M, N376L,
 N376I, or N376F.

30 In an embodiment polypeptide has a substitution is at position corresponding to
 339 and wherein the aminoacid at that position is replaced by an aminoacid that has a
 side chain hydrophobicity of -30 to 10 Δt_R (G,S,N,Q,H,K,R).

In an embodiment polypeptide has a substitution is at position corresponding to
 339 and wherein the aminoacid at that position is replaced by an aminoacid that has a

side chain hydrophobicity of -30 to 10 Δt_R and a van der Waals volume of 60 to 160 $\text{\AA}^3$. (S, N, Q, H, K, R).

In an embodiment polypeptide has a substitution is at position corresponding to 339 and wherein the aminoacid at that position is replaced by an aminoacid that has a
5 side chain hydrophobicity of -30 to 10 Δt_R and a van der Waals volume of 80 to 160 $\text{\AA}^3$. (N, Q, H, K, R).

In an embodiment polypeptide has a substitution is at position corresponding to 339 and wherein the aminoacid at that position is replaced by an aminoacid that has a
10 side chain hydrophobicity of -30 to 10 Δt_R and a van der Waals volume of 100 to 160 $\text{\AA}^3$. (Q, H, K, R).

In an embodiment polypeptide has a substitution is at position corresponding to 339 and wherein the aminoacid at that position is replaced by an aminoacid that has a
side chain hydrophobicity of -30 to 0 Δt_R and a van der Waals volume of 100 to 160 $\text{\AA}^3$. (Q, K, R).

15 In an embodiment polypeptide has a substitution is at position corresponding to 339 and wherein the aminoacid at that position is replaced by an aminoacid that has a side chain hydrophobicity of -30 to 0 Δt_R and a van der Waals volume of 120 to 160 $\text{\AA}^3$. (K, R).

In an embodiment polypeptide has a substitution is at position corresponding to
20 339 and wherein the aminoacid at that position is replaced by an aminoacid that has a side chain hydrophobicity of -30 to 10 Δt_R and a van der Waals volume of 80 to 120 $\text{\AA}^3$. (N, Q, H).

In an embodiment polypeptide has a substitution is at position corresponding to 339 and wherein the aminoacid at that position is replaced by an aminoacid that has a
25 side chain hydrophobicity of -30 to 10 Δt_R and a van der Waals volume of 80 to 120 $\text{\AA}^3$. (N, Q).

In an embodiment, the substitution is M339G, M339S, M339N, M339Q, M339H, M339K, M339R, or M339V.

In an embodiment the polypeptide has one or more aminoacid corresponding to
30 339N/V and/or 376I/M/V.

The permeases belong to the Major Facilitator Superfamily (MFS). This is defined hereinbelow. Cellular transport systems allow the uptake of essential nutrients and ions, and excretion of products of metabolism and deleterious substances. In addition, transport systems play a role in the communication between cells and the environment.

Also, they are an essential part of the cell system to yield or consume energy-supplying molecules, such as ATP.

The transport of solutes by primary active transporters is energy-driven in the first place, such as by energy supplied from ATP hydrolysis, photon absorption, electron flow, substrate decarboxylation, or methyl transfer. If charged molecules are pumped in one direction as a consequence of the consumption of a primary cellular energy source, an electrochemical potential is the result. The resulting chemiosmotic gradient can then be used to drive the transport of additional molecules via secondary carrier structures which just facilitate the transport of one or more molecules across the membrane.

The last two decades the existence of a multitude of previously unknown protein families of primary and secondary transporters has been clarified by the emergence of genomics strategies and making use of the many performed biochemical and molecular genetics studies. The two main transporter families of which proteins were found throughout all living organism are of the ATP-binding cassette (ABC) superfamily and the major facilitator superfamily (MFS), also known as the uniporter-symporter-antiporter family. Whereas ABC family permeases consist of multiple components and are primary active transporters, capable of transporting both small molecules and macromolecules only after generating energy through ATP hydrolysis, the MFS transporters consist of a single polypeptide of a secondary carrier which facilitates transport of small solutes in response to a chemiosmotic ion gradient. ABC superfamily and MFS proteins account for almost half of the solute transporters encoded within the microbe genomes (reviewed by Pao *et al*, 1998, Microbiol Mol Biol Rev.; **62** pp.1–34, and Saier *et al*, 1999, J Mol Microbiol Biotechnol, **1** pp.257-279).

Suitable permease polypeptide sequences can contain one or more of the following motifs:

- a) G-R-x(3)-G-x(3)-G-x(11)-E-x(5)-[LIVM]-R-G-x(12)-[GA] ;
- b) R-x(14)-G-x(2)-Y-x(2)-[YF]-[YF]-[GSAL]; .
- c) V-x(15)-[GNR]-[RH]-R-x(2)-[LM]-x(2)-[GA]

Motif (a) is corresponds to residues 179-221 in Gal2; motif (b) is corresponds to residues 330-353 in Gal2; motif (c) is corresponds to residues 375-399 in Gal2.

In an embodiment the polypeptide comprises a motif G-R-x(3)-G-x(3)-G-x(11)-E-x(5)-[LIVM]-R-G-x(12)-[GA].

The claimed method comprises modeling a permease polypeptide sequence onto the published crystal structure of the xylose- or glucose-bound *Escherichia coli* xylose permease Xyle (respectively, PDB code 4GBY & 4GBZ in the PDB database, <http://www.pdb.org>) to identify the amino acid positions in the channel of the permease that directly interact with the bound sugar (called the first-shell residues in the art), and the residues that interact with the first shell residues (called the second shell residues in the art). Suitable modeling software to construct such models are YASARA, Prime (Schrodinger Inc.) or MODELLER using the default settings. Alternatively, the sugar-specificity-altering first and second shell amino acid positions in a permease polypeptide sequence can be identified by a global pairwise alignment of the permease sequence with the Gal2 sequence SEQ ID NO: 59 using the NEEDLE protocol described below. An example alignment for Gal2 and Hxt's from *Saccharomyces cerevisiae* is given in Figure 10, which shows how alignment can be used to identify the corresponding amino acid positions in the different yeast Hxt's. The amino acid positions herein thus refer to SEQ ID NO: 59 that describes Gal2 or to corresponding amino acid positions in other polypeptides, in particular other permease polypeptides. For example, the corresponding position of the position N376 in Gal2 (SEQ ID NO: 59) in Hxt1 is N370, in Hxt2 N361, in Hxt3 N367, in Hxt4SC N376, in Hxt4RN N376, in Hxt5 N391, in Hxt6/7 N370, in Hxt8 N372, in Hxt9 N366, in Hxt10 N354, in Hxt11 N366, in Hxt12 N256, in Hxt13 N363, in Hxt14 N387, in Hxt15 N366, in Hxt16 N366 and in Hxt17 N363. Similarly, the corresponding position of N346 in Gal2 (SEQ ID NO: 59) in Hxt1 is D340, in Hxt2 N331, in Hxt3 D337, in Hxt4SC D346, in Hxt4RN D346, in Hxt5 D361, in Hxt6/7 D340, in Hxt8 D342, in Hxt9 D336, in Hxt10 C324, in Hxt11 D336, in Hxt12 D226, in Hxt13 E333, in Hxt14 I357, in Hxt15 E336, in Hxt16 E336 and in Hxt17 E333. This can be similarly done for other MFS Superfamily transporters, so that corresponding positions in these polypeptides corresponding to the positions in SEQ ID NO: 59 can be obtained.

A person skilled in the art can subsequently mutate the identified amino acid positions in the permease polypeptide to all other 19 amino acids, and screen for improved C5 sugar uptake and/or reduced C6 sugar uptake of the mutant permease, as described in Example 4 and 5.

For instance, for a polypeptide having a mutation at a position corresponding to one or more position corresponding to N376 of SEQ ID NO: 59, the mutations at the positions corresponding to N376 may be a substitution with C, P, G, A, V, L, I, M, F, W, Y, H, S, T, N, Q, D, E, K, R or a deletion. X may be any amino acid, X(2) means two X.

Herein, Gal2 is a facilitated diffusion transporter required for both the high-affinity galactokinase-dependent and low-affinity galactokinase-independent galactose transport processes. It belongs to the major facilitator superfamily, sugar transporter (TC 2.A.1.1) family. "Permease polypeptide", is also designated herein as "polypeptide permease" or
5 "polypeptide". "Permease polypeptide polynucleotide", is herein a polynucleotide that encodes the permease polypeptide.

In an embodiment of the invention, the permease polypeptide has at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%,
10 at least 95%, at least 96%, at least 97%, at least 98%, at least 99% sequence identity with SEQ ID NO: 59.

Herein mutations are indicated by one letter aminoacids and positions of these amino acids. For example, A6 herein indicates an amino acid (one letter code) at a certain position in SEQ ID NO: 59, here A (Alanine) at position 6 of the protein. A6
15 (L/N/Q/G/V/I/Y/S/E/K) indicates herein mutation of amino acid at a certain position, here A (Alanine) at position 6 of the protein is exchanged for any of L (Leucine), N (Asparagine), Q (Glutamine), G (Glycine), V (Valine), I (Isoleucine), Y (Tyrosine), S (Serine), E (Glutamic acid) or K (Lysine).

In an embodiment, the polypeptide has xylose transport activity.

20 In an embodiment the polypeptide has reduced glucose affinity compared to the polypeptide of SEQ ID NO: 59.

The permease polypeptide of the invention may have one or more alternative and/or additional activities other than that of sugar permease activity.

As set out above, a permease polypeptide of the invention will typically have
25 sugar permease activity. However, a permease polypeptide of the invention may have one or more of the activities set out above in addition to or alternative to that activity.

Polynucleotide sequence

With the permease polypeptide and its amino acid sequence as disclosed herein, the skilled person may determine suitable polynucleotides that encode the permease polypeptide.

5 In an embodiment the polynucleotide is a variant polynucleotide having at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% sequence identity with SEQ ID NO: 56, and encodes the polypeptide as described in claims 1 to
10 11.

The invention therefore provides polynucleotide sequences comprising the gene encoding the permease polypeptide, as well as its coding sequence.

The polynucleotides of the invention may be isolated or synthesized. The permease polypeptides and permease polypeptide polynucleotides herein may be
15 synthetic polypeptides, respectively polynucleotides. The synthetic polynucleotides may be optimized in codon use, preferably according to the methods described in WO2006/077258 and/or PCT/EP2007/055943, which are herein incorporated by reference. PCT/EP2007/055943 addresses codon-pair optimization.

The term refers to a polynucleotide molecule, which is a ribonucleic acid (RNA) or
20 deoxyribonucleic acid (DNA) molecule, either single stranded or double stranded. A polynucleotide may either be present in isolated form, or be comprised in recombinant nucleic acid molecules or vectors, or be comprised in a host cell.

The word "polypeptide" is used herein for chains containing more than seven amino acid residues. All oligopeptide and polypeptide formulas or sequences herein are
25 written from left to right and in the direction from amino terminus to carboxy terminus. The one-letter code of amino acids used herein is commonly known in the art.

By "isolated" polypeptide or protein is intended a polypeptide or protein removed from its native environment. For example, recombinantly produced polypeptides and proteins expressed in host cells are considered isolated for the purpose of the invention
30 as are native or recombinant polypeptides which have been substantially purified by any suitable technique such as, for example, the single-step purification method disclosed in Smith and Johnson, Gene 67:31-40 (1988).

The polynucleotides of the present invention, such as a polynucleotide encoding the permease polypeptide can be isolated or synthesized using standard molecular biology techniques and the sequence information provided herein.

The polynucleotide encoding the permease polypeptide of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis.

Transformation

The polynucleotides according to the invention may be expressed in a suitable host. The invention thus relates to a transformed host cell. In an embodiment, the host cell may be transformed with a nucleic acid construct that comprises a polynucleotide that encodes the polypeptide according to the invention defined before. Therefore standard transformation techniques may be used.

In an embodiment the transformed host cell comprises a heterologous nucleotide that encodes a polypeptide according to claim 1 or encodes a polypeptide having substitution F85Q/V, T89V, V187A/F, I218S, T219A, Q341S/A, N346V, T380A, F444L/V, T448A, T449F, T451G or W455M of sequence ID NO: 59, and in an embodiment thereof the host cell is *Saccharomyces cerevisiae*.

In an embodiment the transformed host is transformed with a polynucleotide that encodes a polypeptide that is a mutant of a polypeptide that is native in the untransformed host cell.

In an embodiment the polypeptide that is native in the untransformed host cell is a member of the the Major Facilitator Superfamily (MFS) transporters, in an embodiment a hexose transporter polypeptide..

In an embodiment the polypeptide that is native in the untransformed host cell is a transporter polypeptide chosen from the list consisting of Gal2, Hxt1, Hxt2, Hxt3, Hxt4, Hxt5, Hxt6, Hxt7, Hxt8, Hxt9, Hxt10, Hxt11, Hxt12, Hxt13, Hxt14, Hxt15, Hxt16 and Hxt17.

The invention further relates to a nucleic acid construct comprising the polynucleotide as described before, e.g. a vector.

Another aspect of the invention thus pertains to vectors, including cloning and expression vectors, comprising a polynucleotide of the invention encoding a permease

polypeptide protein or a functional equivalent thereof and methods of growing, transforming or transfecting such vectors in a suitable host cell, for example under conditions in which expression of a permease of the invention occurs. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another
5 nucleic acid to which it has been linked.

Polynucleotides of the invention can be incorporated into a recombinant replicable vector, for example a cloning or expression vector. The vector may be used to replicate the nucleic acid in a compatible host cell. Thus in a further embodiment, the invention provides a method of making polynucleotides of the invention by introducing a
10 polynucleotide of the invention into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about replication of the vector. The vector may be recovered from the host cell. Suitable host cells are described below.

It will be appreciated by those skilled in the art that the design of the expression
15 vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The vectors, such as expression vectors, of the invention can be introduced into host cells to thereby produce proteins or peptides, encoded by nucleic acids as described herein. The vectors, such as recombinant expression vectors, of the invention can be designed for expression of permease
20 polypeptide proteins in prokaryotic or eukaryotic cells.

For example, permease polypeptides can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors), filamentous fungi, yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA
25 (1990). Representative examples of appropriate hosts are described hereafter.

Appropriate culture mediums and conditions for the above-described host cells are known in the art.

For most filamentous fungi and yeast, the vector or expression construct is preferably integrated in the genome of the host cell in order to obtain stable
30 transformants. However, for certain yeasts also suitable episomal vectors are available into which the expression construct can be incorporated for stable and high level expression, examples thereof include vectors derived from the 2 μ and pKD1 plasmids of *Saccharomyces* and *Kluyveromyces*, respectively, or vectors containing an AMA sequence (e.g. AMA1 from *Aspergillus*). In case the expression constructs are integrated

in the host cells genome, the constructs are either integrated at random loci in the genome, or at predetermined target loci using homologous recombination, in which case the target loci preferably comprise a highly expressed gene.

Accordingly, expression vectors useful in the present invention include
5 chromosomal-, episomal- and virus-derived vectors e.g., vectors derived from bacterial plasmids, bacteriophage, yeast episome, yeast chromosomal elements, viruses such as baculoviruses, papova viruses, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as
10 cosmids and phagemids.

When the polypeptide according to the invention is to be secreted from the host cell into the cultivation medium, an appropriate signal sequence can be added to the polypeptide in order to direct the *de novo* synthesized polypeptide to the secretion route of the host cell. The person skilled in the art knows to select an appropriate signal
15 sequence for a specific host.

The vector may further include sequences flanking the polynucleotide giving rise to RNA which comprise sequences homologous to eukaryotic genomic sequences or viral genomic sequences. This will allow the introduction of the polynucleotides of the invention into the genome of a host cell.

20 An integrative cloning vector may integrate at random or at a predetermined target locus in the chromosome(s) of the host cell into which it is to be integrated.

The vector system may be a single vector, such as a single plasmid, or two or more vectors, such as two or more plasmids, which together contain the total DNA to be introduced into the genome of the host cell.

25 The vector may contain a polynucleotide of the invention oriented in an antisense direction to provide for the production of antisense RNA.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized
30 techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, transduction, infection, lipofection, cationic lipidmediated transfection or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (*Molecular Cloning: A Laboratory Manual*, 2nd, ed. Cold Spring Harbor

Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), Davis et al., *Basic Methods in Molecular Biology* (1986) and other laboratory manuals.

As indicated before, the invention provides an isolated polypeptide having the amino acid sequence according to SEQ ID NO: 59 with the mutations indicated in claim 1.

The permease polypeptide according to the invention can be recovered and purified from recombinant cell cultures by methods known in the art. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

The invention also features biologically active fragments of the polypeptides according to the invention.

Provided also are host cells, comprising a polynucleotide or vector of the invention. The polynucleotide may be heterologous to the genome of the host cell. The term "heterologous", usually with respect to the host cell, means that the polynucleotide does not naturally occur in the genome of the host cell or that the polypeptide is not naturally produced by that cell.

In another embodiment, the invention features cells, e.g., transformed host cells or recombinant host cells that contain a nucleic acid encompassed by the invention. A "transformed cell" or "recombinant cell" is a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a nucleic acid according to the invention. Both prokaryotic and eukaryotic cells are included, e.g., bacteria, fungi, yeast, and the like, especially preferred are yeast cells including e.g. *Saccharomyces*, for example *Saccharomyces cerevisiae*.

A host cell can be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in a specific, desired fashion. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may facilitate optimal functioning of the protein.

Various host cells have characteristic and specific mechanisms for post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems familiar to those of skill in the art of molecular biology and/or microbiology can be chosen to ensure the desired and correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product can be used. Such host cells are well known in the art.

If desired, a cell as described above may be used to in the preparation of a polypeptide according to the invention. Such a method typically comprises cultivating a host cell (e. g. transformed or transfected with an expression vector as described above) under conditions to provide for expression (by the vector) of a coding sequence encoding the polypeptide, and optionally recovering the expressed polypeptide. Polynucleotides of the invention can be incorporated into a recombinant replicable vector, e. g. an expression vector. The vector may be used to replicate the nucleic acid in a compatible host cell. Thus in a further embodiment, the invention provides a method of making a polynucleotide of the invention by introducing a polynucleotide of the invention into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about the replication of the vector. The vector may be recovered from the host cell.

The vectors may be transformed or transfected into a suitable host cell as described above to provide for expression of a polypeptide of the invention. This process may comprise culturing a host cell transformed with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the polypeptide.

Herein standard isolation, hybridization, transformation and cloning techniques are used (e. g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Homology & Identity

Amino acid or nucleotide sequences are said to be homologous when exhibiting a certain level of similarity. Two sequences being homologous indicate a common evolutionary origin. Whether two homologous sequences are closely related or more

distantly related is indicated by “percent identity” or “percent similarity”, which is high or low respectively. Although disputed, to indicate “percent identity” or “percent similarity”, “level of homology” or “percent homology” are frequently used interchangeably.

A comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. The skilled person will be aware of the fact that several different computer programs are available to align two sequences and determine the homology between two sequences (Kruskal, J. B. (1983) An overview of sequence comparison In D. Sankoff and J. B. Kruskal, (ed.), Time warps, string edits and macromolecules: the theory and practice of sequence comparison, pp. 1-44 Addison Wesley). The percent identity between two amino acid sequences can be determined using the Needleman and Wunsch algorithm for the alignment of two sequences. (Needleman, S. B. and Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453). The algorithm aligns amino acid sequences as well as nucleotide sequences. The Needleman-Wunsch algorithm has been implemented in the computer program NEEDLE. For the purpose of this invention the NEEDLE program from the EMBOSS package was used (version 2.8.0 or higher, *EMBOSS: The European Molecular Biology Open Software Suite* (2000) Rice, P. Longden, I. and Bleasby, A. Trends in Genetics 16, (6) pp276—277, <http://emboss.bioinformatics.nl/>). For protein sequences, EBLOSUM62 is used for the substitution matrix. For nucleotide sequences, EDNAFULL is used. Other matrices can be specified. The optional parameters used for alignment of amino acid sequences are a gap-open penalty of 10 and a gap extension penalty of 0.5. The skilled person will appreciate that all these different parameters will yield slightly different results but that the overall percentage identity of two sequences is not significantly altered when using different algorithms.

Global Homology Definition

The homology or identity is the percentage of identical matches between the two full sequences over the total aligned region including any gaps or extensions. The homology or identity between the two aligned sequences is calculated as follows: Number of corresponding positions in the alignment showing an identical amino acid in both sequences divided by the total length of the alignment including the gaps. The identity defined as herein can be obtained from NEEDLE and is labelled in the output of the program as “IDENTITY”.

Longest Identity Definition

The homology or identity between the two aligned sequences is calculated as follows: Number of corresponding positions in the alignment showing an identical amino acid in both sequences divided by the total length of the alignment after subtraction of
5 the total number of gaps in the alignment. The identity defined as herein can be obtained from NEEDLE by using the NOBRIEF option and is labelled in the output of the program as "longest-identity".

The various embodiments of the invention described herein may be cross-
10 combined.

The sugar composition

The sugar composition according to the invention comprises glucose, arabinose and xylose. Any sugar composition may be used in the invention that suffices those
15 criteria. Optional sugars in the sugar composition are galactose and mannose. In a preferred embodiment, the sugar composition is a hydrolysate of one or more lignocellulosic material. Lignocellulose herein includes hemicellulose and hemicellulose parts of biomass. Also lignocellulose includes lignocellulosic fractions of biomass. Suitable lignocellulosic materials may be found in the following list: orchard primings, chaparral, mill waste, urban wood waste, municipal waste, logging waste, forest
20 thinnings, short-rotation woody crops, industrial waste, wheat straw, oat straw, rice straw, barley straw, rye straw, flax straw, soy hulls, rice hulls, rice straw, corn gluten feed, oat hulls, sugar cane, corn stover, corn stalks, corn cobs, corn husks, switch grass, miscanthus, sweet sorghum, canola stems, soybean stems, prairie grass, gamagrass, foxtail; sugar beet pulp, citrus fruit pulp, seed hulls, cellulosic animal wastes, lawn
25 clippings, cotton, seaweed, trees, softwood, hardwood, poplar, pine, shrubs, grasses, wheat, wheat straw, sugar cane bagasse, corn, corn husks, corn hobs, corn kernel, fiber from kernels, products and by-products from wet or dry milling of grains, municipal solid waste, waste paper, yard waste, herbaceous material, agricultural residues, forestry
30 residues, municipal solid waste, waste paper, pulp, paper mill residues, branches, bushes, canes, corn, corn husks, an energy crop, forest, a fruit, a flower, a grain, a grass, a herbaceous crop, a leaf, bark, a needle, a log, a root, a sapling, a shrub, switch grass, a tree, a vegetable, fruit peel, a vine, sugar beet pulp, wheat midlings, oat hulls, hard or soft wood, organic waste material generated from an agricultural process,

forestry wood waste, or a combination of any two or more thereof.

An overview of some suitable sugar compositions derived from lignocellulose and the sugar composition of their hydrolysates is given in table 1. The listed lignocelluloses include: corn cobs, corn fiber, rice hulls, melon shells, sugar beet pulp, wheat straw, sugar cane bagasse, wood, grass and olive pressings.

Table 1: Overview of sugar compositions from lignocellulosic materials. Gal=galactose, Xyl=xylose, Ara=arabinose, Man=mannose, Glu=glucose, Rham=rhamnose. The percentage galactose (% Gal) and literature source is given.

Lignocellulosic material	Gal	Xyl	Ara	Man	Glu	Rham	Sum	% Gal.
Corn cob a	10	286	36		227	11	570	1,7
Corn cob b	131	228	160		144		663	19,8
Rice hulls a	9	122	24	18	234	10	417	2,2
Rice hulls b	8	120	28		209	12	378	2,2
Melon Shells	6	120	11		208	16	361	1,7
Sugar beet pulp	51	17	209	11	211	24	523	9,8
Wheat straw Idaho	15	249	36		396		696	2,2
Corn fiber	36	176	113		372		697	5,2
Cane Bagasse	14	180	24	5	391		614	2,3
Corn stover	19	209	29		370		626	
Athel (wood)	5	118	7	3	493		625	0,7
Eucalyptus (wood)	22	105	8	3	445		583	3,8
CWR (grass)	8	165	33		340		546	1,4
JTW (grass)	7	169	28		311		515	1,3
MSW	4	24	5	20	440		493	0,9
Reed Canary Grass Veg	16	117	30	6	209	1	379	4,2
Reed Canary Grass Seed	13	163	28	6	265	1	476	2,7
Olive pressing residu	15	111	24	8	329		487	3,1

It is clear from table 1 that in these lignocelluloses a high amount of sugar is presence in de form of glucose, xylose, arabinose and galactose. The conversion of

glucose, xylose, arabinose and galactose to fermentation product is thus of great economic importance. Also mannose is present in some lignocellulose materials be it usually in lower amounts than the previously mentioned sugars. Advantageously therefore also mannose is converted by the transformed host cell.

5

The transformed host cell

In an embodiment, the transformed host cell may comprise one or more copies of xylose isomerase gene and/or one or more copies of xylose reductase and/or xylitol dehydrogenase, and two to ten copies of *araA*, *araB* and *araD*, genes, wherein these

10

genes are integrated into the cell genome.

In one embodiment, the transformed host cell comprises genes, for example the above xylose isomerase gene and/or one or more copies of xylose reductase and/or xylitol dehydrogenase, and two to ten copies of *araA*, *araB* and *araD*, genes, are integrated into the transformed host cell genome.

15

The number of copies may be determined by the skilled person by any known method. In the examples, a suitable method is described.

IN an embodiment, the transformed host cell is able to ferment glucose, arabinose, xylose and galactose.

In an embodiment, the cell is capable of converting 90% or more glucose, xylose

20

arabinose, galactose and mannose available, into a fermentation product. In an embodiment, cell is capable of converting 91% or more, 92% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more or 100% of all glucose, xylose arabinose, galactose and mannose available, into a fermentation product.

In one embodiment of the invention the transformed host cell is able to ferment

25

one or more additional sugar, preferably C5 and/or C6 sugar e.g. mannose. In an embodiment of the invention the transformed host cell comprises one or more of: a *xyIA*-gene, *XYL1* gene and *XYL2* gene and/or *XKS1*-gene, to allow the transformed host cell to ferment xylose; deletion of the aldose reductase (*GRE3*) gene; overexpression of PPP-genes *TAL1*, *TKL1*, *RPE1* and *RKI1* to allow the increase of the flux through the

30

pentose phosphate pathway in the cell.

In an embodiment, the transformed host cell is an industrial cell, more preferably an industrial yeast. An industrial cell and industrial yeast cell may be defined as follows. The living environments of (yeast) cells in industrial processes are significantly different from that in the laboratory. Industrial yeast cells must be able to perform well under

multiple environmental conditions which may vary during the process. Such variations include change in nutrient sources, pH, ethanol concentration, temperature, oxygen concentration, etc., which together have potential impact on the cellular growth and ethanol production of *Saccharomyces cerevisiae*. Under adverse industrial conditions, the environmental tolerant strains should allow robust growth and production. Industrial yeast strains are generally more robust towards these changes in environmental conditions which may occur in the applications they are used, such as in the baking industry, brewing industry, wine making and the ethanol industry. In one embodiment, the industrial transformed host cell is constructed on the basis of an industrial host cell, wherein the construction is conducted as described hereinafter. Examples of industrial yeast (*S. cerevisiae*) are Ethanol Red® (Fermentis) Fermiol® (DSM) and Thermosacc® (Lallemand).

In an embodiment the transformed host cell is inhibitor tolerant. Inhibitor tolerance is resistance to inhibiting compounds. The presence and level of inhibitory compounds in lignocellulose may vary widely with variation of feedstock, pretreatment method hydrolysis process. Examples of categories of inhibitors are carboxylic acids, furans and/or phenolic compounds. Examples of carboxylic acids are lactic acid, acetic acid or formic acid. Examples of furans are furfural and hydroxy- methylfurfural. Examples of phenolic compounds are vanillin, syringic acid, ferulic acid and coumaric acid. The typical amounts of inhibitors are for carboxylic acids: several grams per liter, up to 20 grams per liter or more, depending on the feedstock, the pretreatment and the hydrolysis conditions. For furans: several hundreds of milligrams per liter up to several grams per liter, depending on the feedstock, the pretreatment and the hydrolysis conditions.

For phenolics: several tens of milligrams per liter, up to a gram per liter, depending on the feedstock, the pretreatment and the hydrolysis conditions.

The transformed host cells according to the invention may be inhibitor tolerant, i.e. they can withstand common inhibitors at the level that they typically have with common pretreatment and hydrolysis conditions, so that the transformed host cells can find broad application, i.e. it has high applicability for different feedstock, different pretreatment methods and different hydrolysis conditions.

In one embodiment, the industrial transformed host cell is constructed on the basis of an inhibitor tolerant host cell, wherein the construction is conducted as described hereinafter. Inhibitor tolerant host cells may be selected by screening strains

for growth on inhibitors containing materials, such as illustrated in Kadar et al, Appl. Biochem. Biotechnol. (2007), Vol. 136-140, 847-858, wherein an inhibitor tolerant *S. cerevisiae* strain ATCC 26602 was selected.

In an embodiment, the transformed host cell is marker-free. As used herein, the
5 term "marker" refers to a gene encoding a trait or a phenotype which permits the selection of, or the screening for, a host cell containing the marker. Marker-free means that markers are essentially absent in the transformed host cell. Being marker-free is particularly advantageous when antibiotic markers have been used in construction of the transformed host cell and are removed thereafter. Removal of markers may be done
10 using any suitable prior art technique, e.g intramolecular recombination. A suitable method of marker removal is illustrated in the examples.

A transformed host cell may be able to convert plant biomass, celluloses, hemicelluloses, pectins, starch, starch derivatives,, for example into fermentable sugars. Accordingly, a transformed host cell may express one or more enzymes such as a
15 cellulase (an endocellulase or an exocellulase), a hemicellulase (an endo- or exo-xylanase or arabinase) necessary for the conversion of cellulose into glucose monomers and hemicellulose into xylose and arabinose monomers, a pectinase able to convert pectins into glucuronic acid and galacturonic acid or an amylase to convert starch into glucose monomers.

The transformed host cell further may comprise those enzymatic activities
20 required for conversion of pyruvate to a desired fermentation product, such as ethanol, butanol, lactic acid, di-terpene, glycosylated di-terpene, 3 -hydroxy- propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, fumaric acid, malic acid, itaconic acid, an amino acid, 1,3- propane-diol, ethylene, glycerol, a β -lactam antibiotic or a
25 cephalosporin.

In an embodiment, the transformed host cell is a cell that is naturally capable of alcoholic fermentation, preferably, anaerobic alcoholic fermentation. A transformed host cell preferably has a high tolerance to ethanol, a high tolerance to low pH (i.e. capable of growth at a pH lower than about 5, about 4, about 3, or about 2.5) and towards organic
30 and/or a high tolerance to elevated temperatures.

Any of the above characteristics or activities of a transformed host cell may be naturally present in the cell or may be introduced or modified by genetic modification.

Construction of the transformed host cell

According to an embodiment, the genes may be introduced in the host cell by introduction into a host cell:

- a) a cluster consisting of the genes *araA*, *araB* and *araD* under control of a strong constitutive promoter
- b) a cluster consisting of PPP-genes *TAL1*, *TKL1*, *RPE1* and *RKI1*, optionally under control of strong constitutive promoter; and deletion of an aldose reductase gene;
- c) a cluster consisting of a *xyIA*-gene and a *XKS1*-gene under control of strong constitutive promoter;
- d) a construct comprising a *xyIA* gene under control of a strong constitutive promoter, which has the ability to integrate into the genome on multiple loci;

and adaptive evolution to produce the transformed host cell. The above cell may be constructed using recombinant expression techniques.

Recombinant expression

The transformed host cell is a recombinant cell. That is to say, a transformed host cell comprises, or is transformed with or is genetically modified with a nucleotide sequence that does not naturally occur in the cell in question.

Techniques for the recombinant expression of enzymes in a cell, as well as for the additional genetic modifications of a transformed host cell are well known to those skilled in the art. Typically such techniques involve transformation of a cell with nucleic acid construct comprising the relevant sequence. Such methods are, for example, known from standard handbooks, such as Sambrook and Russel (2001) "Molecular Cloning: A Laboratory Manual (3rd edition), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, or F. Ausubel *et al.*, eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987). Methods for transformation and genetic modification of host cells are known from e.g. EP-A- 0635 574, WO 98/46772, WO 99/60102, WO 00/37671, WO90/14423, EP-A-0481008, EP-A-0635574 and US 6,265,186.

Typically, the nucleic acid construct may be a plasmid, for instance a low copy plasmid or a high copy plasmid. The cell according to the present invention may comprise a single or multiple copies of the nucleotide sequence encoding a enzyme, for instance by multiple copies of a nucleotide construct or by use of construct which has multiple copies of the enzyme sequence.

The nucleic acid construct may be maintained episomally and thus comprise a sequence for autonomous replication, such as an autosomal replication sequence sequence. A suitable episomal nucleic acid construct may e.g. be based on the yeast 2 μ or pKD1 plasmids (Gleer *et al.*, 1991, Biotechnology 9: 968-975), or the AMA plasmids (Fierro *et al.*, 1995, Curr Genet. 29:482-489). Alternatively, each nucleic acid construct may be integrated in one or more copies into the genome of the cell. Integration into the cell's genome may occur at random by non-homologous recombination but preferably, the nucleic acid construct may be integrated into the cell's genome by homologous recombination as is well known in the art (see e.g. WO90/14423, EP-A-0481008, EP-A-0635 574 and US 6,265,186).

Most episomal or 2 μ plasmids are relatively unstable in yeast, being lost in approximately 10^{-2} or more cells after each generation. Even under conditions of selective growth, only 60% to 95% of the cells retain the episomal plasmid. The copy number of most episomal plasmids ranges from 20-100 per cell of *cir*⁺ hosts. However, the plasmids are not equally distributed among the cells, and there is a high variance in the copy number per cell in populations. Strains transformed with integrative plasmids are extremely stable, even in the absence of selective pressure. However, plasmid loss can occur at approximately 10^{-3} to 10^{-4} frequencies by homologous recombination between tandemly repeated DNA, leading to looping out of the vector sequence. Preferably, the vector design in the case of stable integration is thus, that upon loss of the selection marker genes (which also occurs by intramolecular, homologous recombination) that looping out of the integrated construct is no longer possible. Preferably the genes are thus stably integrated. Stable integration is herein defined as integration into the genome, wherein looping out of the integrated construct is no longer possible. Preferably selection markers are absent. Typically, the enzyme encoding sequence will be operably linked to one or more nucleic acid sequences, capable of providing for or aiding the transcription and/or translation of the enzyme sequence.

The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. For instance, a promoter or enhancer is operably linked to a coding sequence the said promoter or enhancer affects the transcription of the coding sequence.

As used herein, the term "promoter" refers to a nucleic acid fragment that functions to control the transcription of one or more genes, located upstream with respect to the direction of transcription of the transcription initiation site of the gene, and is structurally

identified by the presence of a binding site for DNA-dependent RNA polymerase, transcription initiation sites and any other DNA sequences known to one of skilled in the art. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is active under
5 environmental or developmental regulation.

The promoter that could be used to achieve the expression of a nucleotide sequence coding for an enzyme according to the present invention, may be not native to the nucleotide sequence coding for the enzyme to be expressed, i.e. a promoter that is heterologous to the nucleotide sequence (coding sequence) to which it is operably
10 linked. The promoter may, however, be homologous, i.e. endogenous, to the host cell.

Promoters are widely available and known to the skilled person. Suitable examples of such promoters include e.g. promoters from glycolytic genes, such as the phosphofructokinase (PFK), triose phosphate isomerase (TPI), glyceraldehyde-3 - phosphate dehydrogenase (GPD, TDH3 or GAPDH), pyruvate kinase (PYK),
15 phosphoglycerate kinase (PGK) promoters from yeasts or filamentous fungi; more details about such promoters from yeast may be found in (WO 93/03159). Other useful promoters are ribosomal protein encoding gene promoters, the lactase gene promoter (LAC4), alcohol dehydrogenase promoters (ADH1, ADH4, and the like), and the enolase promoter (ENO). Other promoters, both constitutive and inducible, and enhancers or
20 upstream activating sequences will be known to those of skill in the art. The promoters used in the host cells of the invention may be modified, if desired, to affect their control characteristics. Suitable promoters in this context include both constitutive and inducible natural promoters as well as engineered promoters, which are well known to the person skilled in the art. Suitable promoters in eukaryotic host cells may be *GAL7*, *GAL10*, or
25 *GAL1*, *CYC1*, *HIS3*, *ADH1*, *PGL*, *PH05*, *GAPDH*, *ADC1*, *TRP1*, *URA3*, *LEU2*, *ENO1*, *TPI1*, and *AOX1*. Other suitable promoters include *PDC1*, *GPD1*, *PGK1*, *TEF1*, and *TDH3*.

In a transformed host cell, the 3'-end of the nucleotide acid sequence encoding enzyme preferably is operably linked to a transcription terminator sequence. Preferably
30 the terminator sequence is operable in a host cell of choice, such as e.g. the yeast species of choice. In any case the choice of the terminator is not critical; it may e.g. be from any yeast gene, although terminators may sometimes work if from a non-yeast, eukaryotic, gene. Usually a nucleotide sequence encoding the enzyme comprises a terminator. Preferably, such terminators are combined with mutations that prevent

nonsense mediated mRNA decay in the host transformed host cell (see for example: Shirley et al., 2002, Genetics 161:1465-1482).

The transcription termination sequence further preferably comprises a polyadenylation signal.

5 Optionally, a selectable marker may be present in a nucleic acid construct suitable for use in the invention. As used herein, the term "marker" refers to a gene encoding a trait or a phenotype which permits the selection of, or the screening for, a host cell containing the marker. The marker gene may be an antibiotic resistance gene whereby the appropriate antibiotic can be used to select for transformed cells from
10 among cells that are not transformed. Examples of suitable antibiotic resistance markers include e.g. dihydrofolate reductase, hygromycin-B-phosphotransferase, 3'-O-phosphotransferase II (kanamycin, neomycin and G418 resistance). Antibiotic resistance markers may be most convenient for the transformation of polyploid host cells. Also non-antibiotic resistance markers may be used, such as auxotrophic markers (*URA3*, *TRP1*,
15 *LEU2*) or the *S. pombe* *TPI* gene (described by Russell P R, 1985, Gene 40: 125-130). In a preferred embodiment the host cells transformed with the nucleic acid constructs are marker gene free. Methods for constructing recombinant marker gene free microbial host cells are disclosed in EP-A-O 635 574 and are based on the use of bidirectional markers such as the *A. nidulans* *amdS* (acetamidase) gene or the yeast *URA3* and *LYS2* genes.
20 Alternatively, a screenable marker such as Green Fluorescent Protein, *lacL*, luciferase, chloramphenicol acetyltransferase, beta-glucuronidase may be incorporated into the nucleic acid constructs of the invention allowing to screen for transformed cells.

Optional further elements that may be present in the nucleic acid constructs suitable for use in the invention include, but are not limited to, one or more leader
25 sequences, enhancers, integration factors, and/or reporter genes, intron sequences, centromeres, telomeres and/or matrix attachment (MAR) sequences. The nucleic acid constructs of the invention may further comprise a sequence for autonomous replication, such as an ARS sequence.

The recombination process may thus be executed with known recombination
30 techniques. Various means are known to those skilled in the art for expression and overexpression of enzymes in a transformed host cell. In particular, an enzyme may be overexpressed by increasing the copy number of the gene coding for the enzyme in the host cell, e.g. by integrating additional copies of the gene in the host cell's genome, by

expressing the gene from an episomal multicopy expression vector or by introducing a episomal expression vector that comprises multiple copies of the gene.

Alternatively, overexpression of enzymes in the host cells of the invention may be achieved by using a promoter that is not native to the sequence coding for the enzyme to be overexpressed, i.e. a promoter that is heterologous to the coding sequence to which it is operably linked. Although the promoter preferably is heterologous to the coding sequence to which it is operably linked, it is also preferred that the promoter is homologous, i.e. endogenous to the host cell. Preferably the heterologous promoter is capable of producing a higher steady state level of the transcript comprising the coding sequence (or is capable of producing more transcript molecules, i.e. mRNA molecules, per unit of time) than is the promoter that is native to the coding sequence. Suitable promoters in this context include both constitutive and inducible natural promoters as well as engineered promoters.

In an embodiment, the transformed host cell is markerfree, which means that no auxotrophic or dominant markers, in particular antibiotic resistance markers, are present in the genome or extra-chromosomally.

The coding sequence used for overexpression of the enzymes mentioned above may preferably be homologous to the host cell. However, coding sequences that are heterologous to the host may be used.

Overexpression of an enzyme, when referring to the production of the enzyme in a genetically modified cell, means that the enzyme is produced at a higher level of specific enzymatic activity as compared to the unmodified host cell under identical conditions. Usually this means that the enzymatically active protein (or proteins in case of multi-subunit enzymes) is produced in greater amounts, or rather at a higher steady state level as compared to the unmodified host cell under identical conditions. Similarly this usually means that the mRNA coding for the enzymatically active protein is produced in greater amounts, or again rather at a higher steady state level as compared to the unmodified host cell under identical conditions. Preferably in a host, an enzyme to be overexpressed is overexpressed by at least a factor of about 1.1, about 1.2, about 1.5, about 2, about 5, about 10 or about 20 as compared to a strain which is genetically identical except for the genetic modification causing the overexpression. It is to be understood that these levels of overexpression may apply to the steady state level of the enzyme's activity, the steady state level of the enzyme's protein as well as to the steady state level of the transcript coding for the enzyme.

Adaptation

Adaptation is the evolutionary process whereby a population becomes better suited (adapted) to its habitat or habitats. This process takes place over several to many
5 generations, and is one of the basic phenomena of biology.

The term adaptation may also refer to a feature which is especially important for an organism's survival. Such adaptations are produced in a variable population by the better suited forms reproducing more successfully, by natural selection.

Changes in environmental conditions alter the outcome of natural selection, affecting the selective benefits of subsequent adaptations that improve an organism's
10 fitness under the new conditions. In the case of an extreme environmental change, the appearance and fixation of beneficial adaptations can be essential for survival. A large number of different factors, such as e.g. nutrient availability, temperature, the availability of oxygen, etcetera, can drive adaptive evolution.

15

Fitness

There is a clear relationship between adaptedness (the degree to which an organism is able to live and reproduce in a given set of habitats) and fitness. Fitness is an estimate and a predictor of the rate of natural selection. By the application of natural
20 selection, the relative frequencies of alternative phenotypes will vary in time, if they are heritable.

Genetic changes

When natural selection acts on the genetic variability of the population, genetic
25 changes are the underlying mechanism. By this means, the population adapts genetically to its circumstances. Genetic changes may result in visible structures, or may adjust the physiological activity of the organism in a way that suits the changed habitat.

It may occur that habitats frequently change. Therefore, it follows that the process of adaptation is never finally complete. In time, it may happen that the
30 environment changes gradually, and the species comes to fit its surroundings better and better. On the other hand, it may happen that changes in the environment occur relatively rapidly, and then the species becomes less and less well adapted. Adaptation is a genetic process, which goes on all the time to some extent, also when the population does not change the habitat or environment.

The adaptive evolution

The transformed host cells may in their preparation be subjected to adaptive evolution. A transformed host cell may be adapted to sugar utilisation by selection of mutants, either spontaneous or induced (e.g. by radiation or chemicals), for growth on the desired sugar, preferably as sole carbon source, and more preferably under anaerobic conditions. Selection of mutants may be performed by techniques including serial transfer of cultures as e.g. described by Kuyper *et al.* (2004, FEMS Yeast Res. 4: 655-664) or by cultivation under selective pressure in a chemostat culture. E.g. in a preferred host cell at least one of the genetic modifications described above, including modifications obtained by selection of mutants, confer to the host cell the ability to grow on the xylose as carbon source, preferably as sole carbon source, and preferably under anaerobic conditions. When XI is used as gene to convert xylose, preferably the cell produce essentially no xylitol, e.g. the xylitol produced is below the detection limit or e.g. less than about 5, about 2, about 1, about 0.5, or about 0.3 % of the carbon consumed on a molar basis.

Adaptive evolution is also described e.g. in Wisselink H.W. et al, Applied and Environmental Microbiology Aug. 2007, p. 4881–4891

In one embodiment of adaptive evolution a regimen consisting of repeated batch cultivation with repeated cycles of consecutive growth in different media is applied, e.g. three media with different compositions (glucose, xylose, and arabinose; xylose and arabinose. See Wisselink et al. (2009) Applied and Environmental Microbiology, Feb. 2009, p. 907–914.

Yeast transformation and genetic stability

Genetic engineering, i.e. transformation of yeast cells with recombinant DNA, became feasible for the first time in 1978 [Beggs, 1978; Hinnen et al., 1978]. Recombinant DNA technology in yeast has established itself since then. A multitude of different vector constructs are available. Generally, these plasmid vectors, called shuttle vectors, contain genetic material derived from *E.coli* vectors consisting of an origin of replication and a selectable marker (often the β lactamase gene, *ampR*), which enable them to be propagated in *E.coli* prior to transformation into yeast cells. Additionally, the shuttle vectors contain a selectable marker for selection in yeast. Markers can be genes encoding enzymes for the synthesis of a particular amino acid or nucleotide, so that cells

carrying the corresponding genomic deletion (or mutation) are complemented for auxotrophy or autotrophy. Alternatively, these vectors contain heterologous dominant resistance markers, which provides recombinant yeast cells (i.e. the cells that have taken up the DNA and express the marker gene) resistance towards certain antibiotics, like g418 (Geneticin), hygromycinB or phleomycin. In addition, these vectors may contain a sequence of (combined) restriction sites (multiple cloning site or MCS) which will allow to clone foreign DNA into these sites, although alternative methods exist as well.

Traditionally, four types of shuttle vectors can be distinguished by the absence or presence of additional genetic elements:

- Integrative plasmids (YIp) which by homologous recombination are integrated into the host genome at the locus of the marker or another gene, when this is opened by restriction and the linearized DNA is used for transformation of the yeast cells. This generally results in the presence of one copy of the foreign DNA inserted at this particular site in the genome.
- Episomal plasmids (YEp) which carry part of the 2 μ plasmid DNA sequence necessary for autonomous replication in yeast cells. Multiple copies of the transformed plasmid are propagated in the yeast cell and maintained as episomes.
- Autonomously replicating plasmids (YRp) which carry a yeast origin of replication (ARS, autonomously replicated sequence) that allows the transformed plasmids to be propagated several hundred-fold.
- CEN plasmids (YCp) which carry in addition to an ARS sequence a centromeric sequence (derived from one of the nuclear chromosomes) which normally guarantees stable mitotic segregation and usually reduces the copy number of self-replicated plasmid to just one.

These plasmids are being introduced into the yeast cells by transformation. Transformation of yeast cells may be achieved by several different techniques, such as permeabilization of cells with lithium acetate (Ito et al, 1983) and electroporation methods.

In commercial application of recombinant microorganisms, plasmid instability is the most important problem. Instability is the tendency of the transformed cells to lose their engineered properties because of changes to, or loss of, plasmids. This issue is discussed in detail by Zhang et al (Plasmid stability in recombinant *Saccharomyces cerevisiae*. Biotechnology Advances, Vol. 14, No. 4, pp. 401-435, 1996). Strains

transformed with integrative plasmids are extremely stable, even in the absence of selective pressure (Sherman, F. http://dbb.urmc.rochester.edu/labs/sherman_f/yeast/9.html and references therein).

5 The heterologous DNA is usually introduced into the organism in the form of extra-chromosomal plasmids (YEp, YCp and YRp). Unfortunately, it has been found with both bacteria and yeasts that the new characteristics may not be retained, especially if the selection pressure is not applied continuously. This is due to the segregational instability of the hybrid plasmid when recombinant cells grow for a long period of time.
10 This leads to population heterogeneity and clonal variability, and eventually to a cell population in which the majority of the cells has lost the properties that were introduced by transformation. If vectors with auxotrophic markers are being used, cultivation in rich media often leads to rapid loss of the vector, since the vector is only retained in minimal media. The alternative, the use of dominant antibiotic resistance markers, is often not
15 compatible with production processes. The use of antibiotics may not be desired from a registration point of view (the possibility that trace amounts of the antibiotic end up in the end product) or for economic reasons (costs of the use of antibiotics at industrial scale).

 Loss of vectors leads to problems in large scale production situations. Alternative methods for introduction of DNA do exist for yeasts, such as the use of integrating
20 plasmids (YIp). The DNA is integrated into the host genome by recombination, resulting in high stability. (Caunt, P. Stability of recombinant plasmids in yeast. Journal of Biotechnology 9(1988) 173 – 192). We have found that an integration method using the host transposons are a good alternative. In an embodiment genes may be integrated into the transformed host cell genome. Initial introduction (i.e. before adaptive evolution) of
25 multiple copies be executed in any way known in the art that leads to introduction of the genes. In an embodiment, this may be accomplished using a vector with parts homologous to repeated sequences (transposons), of the host cell. When the host cell is a yeast cell, suitable repeated sequences are the long terminal repeats (LTR) of the Ty element, known as delta sequence. Ty elements fall into two rather similar subfamilies
30 called Ty1 and Ty2. These elements are about 6 kilobases (kb) in length and are bounded by long terminal repeats (LTR), sequences of about 335 base pairs (Boeke JD et al, The *Saccharomyces cerevisiae* Genome Contains Functional and Nonfunctional Copies of Transposon Ty1. Molecular and Cellular Biology, Apr. 1988, p. 1432-1442 Vol. 8, No. 4). In the fully sequenced *S. cerevisiae* strain, S288c, the most abundant

transposons are Ty1 (31 copies) and Ty2 (13 copies) (Gabriel A, Dapprich J, Kunkel M, Gresham D, Pratt SC, et al. (2006) Global mapping of transposon location. PLoS Genet 2(12): e212.doi:10.1371/journal.pgen.0020212). These transposons consist of two overlapping open reading frames (ORFs), each of which encode several proteins. The coding regions are flanked by the aforementioned, nearly identical LTRs. Other, but less abundant and more distinct Ty elements in *S. cerevisiae* comprise Ty3, Ty4 and Ty5. For each family of full-length Ty elements there are an order of magnitude more solo LTR elements dispersed through the genome. These are thought to arise by LTR–LTR recombination of full-length elements, with looping out of the internal protein encoding regions.

The retrotransposition mechanism of the Ty retrotransposon has been exploited to integrate multiple copies throughout the genome (Boeke et al., 1988; Jacobs et al., 1988). The long terminal repeats (LTR) of the Ty element, known as delta sequences, are also good targets for integration by homologous recombination as they exist in about 150-200 copies that are either Ty associated or solo sites (Boeke, 1989; Kingsman and Kingsman, 1988). (Parekh R.N. (1996). An Integrating Vector for Tunable, High Copy, Stable Integration into the Dispersed Ty DELTA Sites of *Saccharomyces cerevisiae*. Biotechnol. Prog. 1996, 12, 16-21). By adaptive evolution, the number of copies may change.

The host cell

The host cell may be any host cell suitable for production of a useful product. A host cell may be any suitable cell, such as a prokaryotic cell, such as a bacterium, or a eukaryotic cell. Typically, the cell will be a eukaryotic cell, for example a yeast or a filamentous fungus.

Yeasts are herein defined as eukaryotic microorganisms and include all species of the subdivision Eumycotina (Alexopoulos, C. J., 1962, In : Introductory Mycology, John Wiley & Sons, Inc. , New York) that predominantly grow in unicellular form.

Yeasts may either grow by budding of a unicellular thallus or may grow by fission of the organism. A preferred yeast as a transformed host cell may belong to the genera *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces* or *Yarrowia*. Preferably the yeast is one capable of anaerobic fermentation, more preferably one capable of anaerobic alcoholic fermentation.

Filamentous fungi are herein defined as eukaryotic microorganisms that include all filamentous forms of the subdivision Eumycotina. These fungi are characterized by a vegetative mycelium composed of chitin, cellulose, and other complex polysaccharides. The filamentous fungi of the suitable for use as a cell of the present invention are morphologically, physiologically, and genetically distinct from yeasts. Filamentous fungal cells may be advantageously used since most fungi do not require sterile conditions for propagation and are insensitive to bacteriophage infections. Vegetative growth by filamentous fungi is by hyphal elongation and carbon catabolism of most filamentous fungi is obligately aerobic. Preferred filamentous fungi as a host cell may belong to the genus *Aspergillus*, *Trichoderma*, *Humicola*, *Acremonium*, *Fusarium* or *Penicillium*. More preferably, the filamentous fungal cell may be a *Aspergillus niger*, *Aspergillus oryzae*, a *Penicillium chrysogenum*, or *Rhizopus oryzae* cell.

In one embodiment the host cell may be yeast.

Preferably the host is an industrial host, more preferably an industrial yeast. An industrial host and industrial yeast cell may be defined as follows. The living environments of yeast cells in industrial processes are significantly different from that in the laboratory. Industrial yeast cells must be able to perform well under multiple environmental conditions which may vary during the process. Such variations include change in nutrient sources, pH, ethanol concentration, temperature, oxygen concentration, etc., which together have potential impact on the cellular growth and ethanol production of *Saccharomyces cerevisiae*. Under adverse industrial conditions, the environmental tolerant strains should allow robust growth and production. Industrial yeast strains are generally more robust towards these changes in environmental conditions which may occur in the applications they are used, such as in the baking industry, brewing industry, wine making and the ethanol industry. Examples of industrial yeast (*S. cerevisiae*) are Ethanol Red® (Fermentis) Fermiol® (DSM) and Thermosacc® (Lallemand).

In an embodiment the host is inhibitor tolerant. Inhibitor tolerant host cells may be selected by screening strains for growth on inhibitors containing materials, such as illustrated in Kadar et al, Appl. Biochem. Biotechnol. (2007), Vol. 136-140, 847-858, wherein an inhibitor tolerant *S. cerevisiae* strain ATCC 26602 was selected.

araA, araB and araD genes

A transformed host cell is capable of using arabinose. A transformed host cell is therefore, be capable of converting L-arabinose into L-ribulose and/or xylulose 5-phosphate and/or into a desired fermentation product, for example one of those mentioned herein.

Organisms, for example *S. cerevisiae* strains, able to produce ethanol from L-arabinose may be produced by modifying a cell introducing the *araA* (L-arabinose isomerase), *araB* (L-ribulokinase) and *araD* (L-ribulose-5-P4-epimerase) genes from a suitable source. Such genes may be introduced into a transformed host cell in order that it is capable of using arabinose. Such an approach is given is described in WO2003/095627. *araA*, *araB* and *araD* genes from *Lactobacillus plantarum* may be used and are disclosed in WO2008/041840. The *araA* gene from *Bacillus subtilis* and the *araB* and *araD* genes from *Escherichia coli* may be used and are disclosed in EP1499708. In another embodiment, *araA*, *araB* and *araD* genes may be derived from at least one of the genus *Clavibacter*, *Arthrobacter* and/or *Gramella*, in particular one of *Clavibacter michiganensis*, *Arthrobacter aurescens*, and/or *Gramella forsetii*, as disclosed in WO 2009011591.

PPP-genes

A transformed host cell may comprise one or more genetic modifications that increases the flux of the pentose phosphate pathway. In particular, the genetic modification(s) may lead to an increased flux through the non-oxidative part of the pentose phosphate pathway. A genetic modification that causes an increased flux of the non-oxidative part of the pentose phosphate pathway is herein understood to mean a modification that increases the flux by at least a factor of about 1.1, about 1.2, about 1.5, about 2, about 5, about 10 or about 20 as compared to the flux in a strain which is genetically identical except for the genetic modification causing the increased flux. The flux of the non-oxidative part of the pentose phosphate pathway may be measured by growing the modified host on xylose as sole carbon source, determining the specific xylose consumption rate and subtracting the specific xylitol production rate from the specific xylose consumption rate, if any xylitol is produced. However, the flux of the non-oxidative part of the pentose phosphate pathway is proportional with the growth rate on xylose as sole carbon source, preferably with the anaerobic growth rate on xylose as sole carbon source. There is a linear relation between the growth rate on xylose as sole

carbon source ($\mu_{\max}$) and the flux of the non-oxidative part of the pentose phosphate pathway. The specific xylose consumption rate (Q_s) is equal to the growth rate (μ) divided by the yield of biomass on sugar (Y_{xs}) because the yield of biomass on sugar is constant (under a given set of conditions: anaerobic, growth medium, pH, genetic background of the strain, etc.; i.e. $Q_s = \mu / Y_{xs}$). Therefore the increased flux of the non-oxidative part of the pentose phosphate pathway may be deduced from the increase in maximum growth rate under these conditions unless transport (uptake is limiting).

One or more genetic modifications that increase the flux of the pentose phosphate pathway may be introduced in the host cell in various ways. These including e.g. achieving higher steady state activity levels of xylulose kinase and/or one or more of the enzymes of the non-oxidative part pentose phosphate pathway and/or a reduced steady state level of unspecific aldose reductase activity. These changes in steady state activity levels may be effected by selection of mutants (spontaneous or induced by chemicals or radiation) and/or by recombinant DNA technology e.g. by overexpression or inactivation, respectively, of genes encoding the enzymes or factors regulating these genes.

In a preferred host cell, the genetic modification comprises overexpression of at least one enzyme of the (non-oxidative part) pentose phosphate pathway. Preferably the enzyme is selected from the group consisting of the enzymes encoding for ribulose-5-phosphate isomerase, ribulose-5-phosphate epimerase, transketolase and transaldolase. Various combinations of enzymes of the (non-oxidative part) pentose phosphate pathway may be overexpressed. E.g. the enzymes that are overexpressed may be at least the enzymes ribulose-5-phosphate isomerase and ribulose-5-phosphate epimerase; or at least the enzymes ribulose-5-phosphate isomerase and transketolase; or at least the enzymes ribulose-5-phosphate isomerase and transaldolase; or at least the enzymes ribulose-5-phosphate epimerase and transketolase; or at least the enzymes ribulose-5-phosphate epimerase and transaldolase; or at least the enzymes transketolase and transaldolase; or at least the enzymes ribulose-5-phosphate epimerase, transketolase and transaldolase; or at least the enzymes ribulose-5-phosphate isomerase, transketolase and transaldolase; or at least the enzymes ribulose-5-phosphate isomerase, ribulose-5-phosphate epimerase, and transaldolase; or at least the enzymes ribulose-5-phosphate isomerase, ribulose-5-phosphate epimerase, and transketolase. In one embodiment of the invention each of the enzymes ribulose-5-phosphate isomerase, ribulose-5-phosphate epimerase, transketolase and transaldolase are overexpressed in the host cell. More preferred is a host cell in which the genetic

modification comprises at least overexpression of both the enzymes transketolase and transaldolase as such a host cell is already capable of anaerobic growth on xylose. In fact, under some conditions host cells overexpressing only the transketolase and the transaldolase already have the same anaerobic growth rate on xylose as do host cells that overexpress all four of the enzymes, i.e. the ribulose-5-phosphate isomerase, ribulose-5-phosphate epimerase, transketolase and transaldolase. Moreover, host cells overexpressing both of the enzymes ribulose-5-phosphate isomerase and ribulose-5-phosphate epimerase are preferred over host cells overexpressing only the isomerase or only the epimerase as overexpression of only one of these enzymes may produce metabolic imbalances.

The enzyme "ribulose 5-phosphate epimerase" (EC 5.1.3.1) is herein defined as an enzyme that catalyses the epimerisation of D-xylulose 5-phosphate into D-ribulose 5-phosphate and vice versa. The enzyme is also known as phosphoribulose epimerase; erythrose-4-phosphate isomerase; phosphoketopentose 3-epimerase; xylulose phosphate 3-epimerase; phosphoketopentose epimerase; ribulose 5-phosphate 3-epimerase; D-ribulose phosphate-3-epimerase; D-ribulose 5-phosphate epimerase; D-ribulose-5-P 3-epimerase; D-xylulose-5-phosphate 3-epimerase; pentose-5-phosphate 3-epimerase; or D-ribulose-5-phosphate 3-epimerase. A ribulose 5-phosphate epimerase may be further defined by its amino acid sequence. Likewise a ribulose 5-phosphate epimerase may be defined by a nucleotide sequence encoding the enzyme as well as by a nucleotide sequence hybridising to a reference nucleotide sequence encoding a ribulose 5-phosphate epimerase. The nucleotide sequence encoding for ribulose 5-phosphate epimerase is herein designated *RPE1*.

The enzyme "ribulose 5-phosphate isomerase" (EC 5.3.1.6) is herein defined as an enzyme that catalyses direct isomerisation of D-ribose 5-phosphate into D-ribulose 5-phosphate and vice versa. The enzyme is also known as phosphopentosisomerase; phosphoriboisomerase; ribose phosphate isomerase; 5-phosphoribose isomerase; D-ribose 5-phosphate isomerase; D-ribose-5-phosphate ketol-isomerase; or D-ribose-5-phosphate aldose-ketose-isomerase. A ribulose 5-phosphate isomerase may be further defined by its amino acid sequence. Likewise a ribulose 5-phosphate isomerase may be defined by a nucleotide sequence encoding the enzyme as well as by a nucleotide sequence hybridising to a reference nucleotide sequence encoding a ribulose 5-phosphate isomerase. The nucleotide sequence encoding for ribulose 5-phosphate isomerase is herein designated *RKI1*.

The enzyme "transketolase" (EC 2.2.1.1) is herein defined as an enzyme that catalyses the reaction: D-ribose 5-phosphate + D-xylulose 5-phosphate $\leftrightarrow$ sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate and vice versa. The enzyme is also known as glycolaldehydetransferase or sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate glycolaldehydetransferase. A transketolase may be further defined by its amino acid. Likewise a transketolase may be defined by a nucleotide sequence encoding the enzyme as well as by a nucleotide sequence hybridising to a reference nucleotide sequence encoding a transketolase. The nucleotide sequence encoding for transketolase is herein designated *TKL1*.

The enzyme "transaldolase" (EC 2.2.1.2) is herein defined as an enzyme that catalyses the reaction: sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate $\leftrightarrow$ D-erythrose 4-phosphate + D-fructose 6-phosphate and vice versa. The enzyme is also known as dihydroxyacetone transferase; dihydroxyacetone synthase; formaldehyde transketolase; or sedoheptulose-7-phosphate : D-glyceraldehyde-3-phosphate glyceronetransferase. A transaldolase may be further defined by its amino acid sequence. Likewise a transaldolase may be defined by a nucleotide sequence encoding the enzyme as well as by a nucleotide sequence hybridising to a reference nucleotide sequence encoding a transaldolase. The nucleotide sequence encoding for transketolase from is herein designated *TAL1*.

Xylose Isomerase or xylose reductase genes

According to the invention, one or more copies of one or more xylose isomerase gene and/or one or more xylose reductase and xylitol dehydrogenase are introduced into the genome of the host cell. The presence of these genetic elements confers on the cell the ability to convert xylose by isomerisation or reduction.

In one embodiment, the one or more copies of one or more xylose isomerase gene are introduced into the genome of the host cell.

A "xylose isomerase" (EC 5.3.1.5) is herein defined as an enzyme that catalyses the direct isomerisation of D-xylose into D-xylulose and/or *vice versa*. The enzyme is also known as a D-xylose ketoisomerase. A xylose isomerase herein may also be capable of catalysing the conversion between D-glucose and D-fructose (and accordingly may therefore be referred to as a glucose isomerase). A xylose isomerase herein may require a bivalent cation, such as magnesium, manganese or cobalt as a cofactor.

Accordingly, such a transformed host cell is capable of isomerising xylose to xylulose. The ability of isomerising xylose to xylulose is conferred on the host cell by transformation of the host cell with a nucleic acid construct comprising a nucleotide sequence encoding a defined xylose isomerase. A transformed host cell isomerises xylose into xylulose by the direct isomerisation of xylose to xylulose.

A unit (U) of xylose isomerase activity may herein be defined as the amount of enzyme producing 1 nmol of xylulose per minute, under conditions as described by Kuyper *et al.* (2003, FEMS Yeast Res. 4: 69-78).

The Xylose isomerase gene may have various origin, such as for example *Piromyces* sp. as disclosed in WO2006/009434. Other suitable origins are *Bacteroides*, in particular *Bacteroides uniformis* as described in PCT/EP2009/52623, *Bacillus*, in particular *Bacillus stearothermophilus* as described in PCT/EP2009/052625.

In another embodiment, one or more copies of one or more xylose reductase and xylitol dehydrogenase genes are introduced into the genome of the host cell. In this embodiment the conversion of xylose is conducted in a two step conversion of xylose into xylulose via a xylitol intermediate as catalysed by xylose reductase and xylitol dehydrogenase, respectively. In an embodiment thereof xylose reductase (XR), xylitol dehydrogenase (XDH), and xylokinase (XK) may be overexpressed, and optionally one or more of genes encoding NADPH producing enzymes are up-regulated and one or more of the genes encoding NADH consuming enzymes are up-regulated, as disclosed in WO 2004085627.

XKS1 gene

A transformed host cell may comprise one or more genetic modifications that increase the specific xylulose kinase activity. Preferably the genetic modification or modifications causes overexpression of a xylulose kinase, e.g. by overexpression of a nucleotide sequence encoding a xylulose kinase. The gene encoding the xylulose kinase may be endogenous to the host cell or may be a xylulose kinase that is heterologous to the host cell. A nucleotide sequence used for overexpression of xylulose kinase in the host cell is a nucleotide sequence encoding a polypeptide with xylulose kinase activity.

The enzyme "xylulose kinase" (EC 2.7.1.17) is herein defined as an enzyme that catalyses the reaction $\text{ATP} + \text{D-xylulose} = \text{ADP} + \text{D-xylulose 5-phosphate}$. The enzyme is also known as a phosphorylating xylulokinase, D-xylulokinase or ATP :D- xylulose 5-phosphotransferase. A xylulose kinase of the invention may be further defined by its

amino acid sequence. Likewise a xylulose kinase may be defined by a nucleotide sequence encoding the enzyme as well as by a nucleotide sequence hybridising to a reference nucleotide sequence encoding a xylulose kinase.

In a transformed host cell, a genetic modification or modifications that increase(s)
5 the specific xylulose kinase activity may be combined with any of the modifications increasing the flux of the pentose phosphate pathway as described above. This is not, however, essential.

Thus, a host cell may comprise only a genetic modification or modifications that increase the specific xylulose kinase activity. The various means available in the art for
10 achieving and analysing overexpression of a xylulose kinase in the host cells of the invention are the same as described above for enzymes of the pentose phosphate pathway. Preferably in the host cells of the invention, a xylulose kinase to be overexpressed is overexpressed by at least a factor of about 1.1, about 1.2, about 1.5, about 2, about 5, about 10 or about 20 as compared to a strain which is genetically
15 identical except for the genetic modification(s) causing the overexpression. It is to be understood that these levels of overexpression may apply to the steady state level of the enzyme's activity, the steady state level of the enzyme's protein as well as to the steady state level of the transcript coding for the enzyme.

20 Aldose reductase (*GRE3*) gene deletion

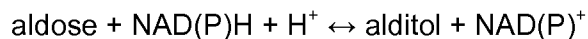
In the embodiment, where XI is used as gene to convert xylose, it may be advantageous to reduce aldose reductase activity. A transformed host cell may therefore comprise one or more genetic modifications that reduce unspecific aldose reductase activity in the host cell. Preferably, unspecific aldose reductase activity is
25 reduced in the host cell by one or more genetic modifications that reduce the expression of or inactivates a gene encoding an unspecific aldose reductase. Preferably, the genetic modification(s) reduce or inactivate the expression of each endogenous copy of a gene encoding an unspecific aldose reductase in the host cell (herein called *GRE3* deletion). Transformed host cells may comprise multiple copies of genes encoding unspecific
30 aldose reductases as a result of di-, poly- or aneu-ploidy, and/or the host cell may contain several different (iso)enzymes with aldose reductase activity that differ in amino acid sequence and that are each encoded by a different gene. Also in such instances preferably the expression of each gene that encodes an unspecific aldose reductase is reduced or inactivated. Preferably, the gene is inactivated by deletion of at least part of

the gene or by disruption of the gene, whereby in this context the term gene also includes any non-coding sequence up- or down-stream of the coding sequence, the (partial) deletion or inactivation of which results in a reduction of expression of unspecific aldose reductase activity in the host cell.

5 A nucleotide sequence encoding an aldose reductase whose activity is to be reduced in the host cell is a nucleotide sequence encoding a polypeptide with aldose reductase activity.

Thus, a host cell comprising only a genetic modification or modifications that reduce(s) unspecific aldose reductase activity in the host cell is specifically included in
10 the invention.

The enzyme "aldose reductase" (EC 1.1.1.21) is herein defined as any enzyme that is capable of reducing xylose or xylulose to xylitol. In the context of the present invention an aldose reductase may be any unspecific aldose reductase that is native (endogenous) to a host cell of the invention and that is capable of reducing xylose
15 or xylulose to xylitol. Unspecific aldose reductases catalyse the reaction:



The enzyme has a wide specificity and is also known as aldose reductase; polyol dehydrogenase (NADP⁺); alditol:NADP oxidoreductase; alditol:NADP⁺ 1-oxidoreductase; NADPH-aldopentose reductase; or NADPH-aldose reductase.

20 A particular example of such an unspecific aldose reductase that is endogenous to *S. cerevisiae* and that is encoded by the GRE3 gene (Traff *et al.*, 2001, Appl. Environ. Microbiol. 67: 5668-74). Thus, an aldose reductase of the invention may be further defined by its amino acid sequence. Likewise an aldose reductase may be defined by the nucleotide sequences encoding the enzyme as well as by a nucleotide sequence
25 hybridising to a reference nucleotide sequence encoding an aldose reductase.

Bioproducts production

Over the years suggestions have been made for the introduction of various organisms for the production of bio-ethanol from crop sugars. In practice, however, all
30 major bio-ethanol production processes have continued to use the yeasts of the genus *Saccharomyces* as ethanol producer. This is due to the many attractive features of *Saccharomyces* species for industrial processes, i. e. , a high acid-, ethanol-and osmo-tolerance, capability of anaerobic growth, and of course its high alcoholic fermentative

capacity. Preferred yeast species as host cells include *S. cerevisiae*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus* or *K fragilis*.

A transformed host cell may be a cell suitable for the production of ethanol. A transformed host cell may, however, be suitable for the production of fermentation
5 products other than ethanol

Such non-ethanolic fermentation products include in principle any bulk or fine chemical that is producible by a eukaryotic microorganism such as a yeast or a filamentous fungus.

A transformed host cell that may be used for production of non-ethanolic
10 fermentation products is a host cell that contains a genetic modification that results in decreased alcohol dehydrogenase activity.

In an embodiment the transformed host cell may be used in a process wherein sugars originating from lignocellulose are converted into ethanol.

15 Lignocellulose

Lignocellulose, which may be considered as a potential renewable feedstock, generally comprises the polysaccharides cellulose (glucans) and hemicelluloses (xylans, heteroxylans and xyloglucans). In addition, some hemicellulose may be present as glucomannans, for example in wood-derived feedstocks. The enzymatic hydrolysis of
20 these polysaccharides to soluble sugars, including both monomers and multimers, for example glucose, cellobiose, xylose, arabinose, galactose, fructose, mannose, rhamnose, ribose, galacturonic acid, glucuronic acid and other hexoses and pentoses occurs under the action of different enzymes acting in concert.

In addition, pectins and other pectic substances such as arabinans may make
25 up considerably proportion of the dry mass of typically cell walls from non-woody plant tissues (about a quarter to half of dry mass may be pectins).

Pretreatment

Before enzymatic treatment, the lignocellulosic material may be pretreated. The
30 pretreatment may comprise exposing the lignocellulosic material to an acid, a base, a solvent, heat, a peroxide, ozone, mechanical shredding, grinding, milling or rapid depressurization, or a combination of any two or more thereof. This chemical pretreatment is often combined with heat-pretreatment, e.g. between 150-220 °C for 1 to 30 minutes.

Enzymatic hydrolysis

The pretreated material is commonly subjected to enzymatic hydrolysis to release sugars that may be fermented according to the invention. This may be executed with conventional methods, e.g. contacting with cellulases, for instance cellobiohydrolase(s), endoglucanase(s), beta-glucosidase(s) and optionally other enzymes. The conversion with the cellulases may be executed at ambient temperatures or at higher temperatures, at a reaction time to release sufficient amounts of sugar(s). The result of the enzymatic hydrolysis is hydrolysis product comprising C5/C6 sugars, herein designated as the sugar composition.

Fermentation

The fermentation process may be an aerobic or an anaerobic fermentation process. An anaerobic fermentation process is herein defined as a fermentation process run in the absence of oxygen or in which substantially no oxygen is consumed, preferably less than about 5, about 2.5 or about 1 mmol/L/h, more preferably 0 mmol/L/h is consumed (i.e. oxygen consumption is not detectable), and wherein organic molecules serve as both electron donor and electron acceptors. In the absence of oxygen, NADH produced in glycolysis and biomass formation, cannot be oxidised by oxidative phosphorylation. To solve this problem many microorganisms use pyruvate or one of its derivatives as an electron and hydrogen acceptor thereby regenerating NAD⁺.

Thus, in a preferred anaerobic fermentation process pyruvate is used as an electron (and hydrogen acceptor) and is reduced to fermentation products such as ethanol, butanol, lactic acid, di-terpene, glycosylated di-terpene, 3 -hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, malic acid, fumaric acid, an amino acid, 1,3-propane-diol, ethylene, glycerol, a β -lactam antibiotic and a cephalosporin.

The fermentation process is preferably run at a temperature that is optimal for the cell. Thus, for most yeasts or fungal host cells, the fermentation process is performed at a temperature which is less than about 42°C, preferably less than about 38°C. For yeast or filamentous fungal host cells, the fermentation process is preferably performed at a temperature which is lower than about 35, about 33, about 30 or about 28°C and at a temperature which is higher than about 20, about 22, or about 25°C.

The ethanol yield on xylose and/or glucose in the process preferably is at least about 50, about 60, about 70, about 80, about 90, about 95 or about 98%. The ethanol yield is herein defined as a percentage of the theoretical maximum yield.

The invention also relates to a process for producing a fermentation product.

5 The fermentation process according to the present invention may be run under aerobic and anaerobic conditions. In an embodiment, the process is carried out under micro-aerophilic or oxygen limited conditions.

10 An anaerobic fermentation process is herein defined as a fermentation process run in the absence of oxygen or in which substantially no oxygen is consumed, preferably less than about 5, about 2.5 or about 1 mmol/L/h, and wherein organic molecules serve as both electron donor and electron acceptors.

15 An oxygen-limited fermentation process is a process in which the oxygen consumption is limited by the oxygen transfer from the gas to the liquid. The degree of oxygen limitation is determined by the amount and composition of the ingoing gasflow as well as the actual mixing/mass transfer properties of the fermentation equipment used. Preferably, in a process under oxygen-limited conditions, the rate of oxygen consumption is at least about 5.5, more preferably at least about 6, such as at least 7 mmol/L/h. A process of the invention may comprise recovery of the fermentation product.

20 In a preferred process the cell ferments both the xylose and glucose, preferably simultaneously in which case preferably a cell is used which is insensitive to glucose repression to prevent diauxic growth. In addition to a source of xylose (and glucose) as carbon source, the fermentation medium will further comprise the appropriate ingredient required for growth of the cell. Compositions of fermentation media for growth of microorganisms such as yeasts are well known in the art

25 The fermentation processes may be carried out in batch, fed-batch or continuous mode. A separate hydrolysis and fermentation (SHF) process or a simultaneous saccharification and fermentation (SSF) process may also be applied. A combination of these fermentation process modes may also be possible for optimal productivity. These processes are described hereafter in more detail.

30

SSF mode

For Simultaneous Saccharification and Fermentation (SSF) mode, the reaction time for liquefaction/hydrolysis or presaccharification step is dependent on the time to realize a desired yield, i.e. cellulose to glucose conversion yield. Such yield is preferably

as high as possible, preferably 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, even 99.5% or more or 99.9% or more.

According to the invention very high sugar concentrations in SHF mode and very high product concentrations (e.g. ethanol) in SSF mode are realized. In SHF operation the glucose concentration is 25g/L or more, 30 g/L or more, 35g/L or more, 40 g/L or more, 45 g/L or more, 50 g/L or more, 55 g/L or more, 60 g/L or more, 65 g/L or more, 70 g/L or more, 75 g/L or more, 80 g/L or more, 85 g/L or more, 90 g/L or more, 95 g/L or more, 100 g/L or more, 110 g/L or more, 120g/L or more or may e.g. be 25g/L-250 g/L, 30g/L-200g/L, 40g/L-200 g/L, 50g/L-200g/L, 60g/L-200g/L, 70g/L-200g/L, 80g/L-200g/L, 90 g/L, 80g/L-200g/L.

Product concentration in SSF mode

In SSF operation, the product concentration (g/L) is dependent on the amount of glucose produced, but this is not visible since sugars are converted to product in the SSF, and product concentrations can be related to underlying glucose concentration by multiplication with the theoretical maximum yield (Yps max in gr product per gram glucose)

The theoretical maximum yield (Yps max in gr product per gram glucose) of a fermentation product can be derived from textbook biochemistry. For ethanol, 1 mole of glucose (180 gr) yields according to normal glycolysis fermentation pathway in yeast 2 moles of ethanol ($=2 \times 46 = 92$ gr ethanol). The theoretical maximum yield of ethanol on glucose is therefore $92/180 = 0.511$ gr ethanol/gr glucose.

For Butanol (MW 74 gr/mole) or iso butanol, the theoretical maximum yield is 1 mole of butanol per mole of glucose. So Yps max for (iso-)butanol = $74/180 = 0.411$ gr (iso-)butanol/gr glucose.

For lactic acid the fermentation yield for homolactic fermentation is 2 moles of lactic acid (MW = 90 gr/mole) per mole of glucose. According to this stoichiometry, the Yps max = 1 gr lactic acid/gr glucose.

For other fermentation products a similar calculation may be made.

SSF mode

In SSF operation the product concentration is 25g * Yps g/L /L or more, 30 * Yps g/L or more, 35g * Yps /L or more, 40 * Yps g/L or more, 45 * Yps g/L or more, 50 * Yps g/L or more, 55 * Yps g/L or more, 60 * Yps g/L or more, 65 * Yps g/L or more, 70 * Yps g/L or more, 75 * Yps g/L or more, 80 * Yps g/L or more, 85 * Yps g/L or more, 90 * Yps g/L or more, 95 * Yps g/L or more, 100 * Yps g/L or more, 110 * Yps g/L or more, 120g/L * Yps or more or may e.g. be 25 * Yps g/L-250 * Yps g/L, 30 * Yps g/L-200 * Yps g/L, 40 * Yps g/L-200 * Yps g/L, 50 * Yps g/L-200 * Yps g/L, 60 * Yps g/L-200 * Yps g/L, 70 * Yps g/L-200 * Yps g/L, 80 * Yps g/L-200 * Yps g/L, 90 * Yps g/L, 80 * Yps g/L-200 * Yps g/L

Accordingly, the invention provides a method for the preparation of a fermentation product, which method comprises:

- a. degrading lignocellulose using a method as described herein; and
 - b. fermenting the resulting material,
- thereby to prepare a fermentation product.

Fermentation product

The fermentation product of the invention may be any useful product. In one embodiment, it is a product selected from the group consisting of ethanol, n-butanol, isobutanol, lactic acid, di-terpene, glycosylated di-terpene, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, fumaric acid, malic acid, itaconic acid, maleic acid, citric acid, adipic acid, an amino acid, such as lysine, methionine, tryptophan, threonine, and aspartic acid, 1,3-propane-diol, ethylene, glycerol, a β -lactam antibiotic and a cephalosporin, vitamins, pharmaceuticals, animal feed supplements, specialty chemicals, chemical feedstocks, plastics, solvents, fuels, including biofuels and biogas or organic polymers, and an industrial enzyme, such as a protease, a cellulase, an amylase, a glucanase, a lactase, a lipase, a lyase, an oxidoreductases, a transferase or a xylanase. For example the fermentation products may be produced by cells according to the invention, following prior art cell preparation methods and fermentation processes, which examples however should herein not be construed as limiting. n-butanol may be

produced by cells as described in WO2008121701 or WO2008086124; lactic acid as described in US2011053231 or US2010137551; 3-hydroxy-propionic acid as described in WO2010010291; acrylic acid as described in WO2009153047.

5 Recovery of the fermentation product

For the recovery of the fermentation product existing technologies are used. For different fermentation products different recovery processes are appropriate. Existing methods of recovering ethanol from aqueous mixtures commonly use fractionation and adsorption techniques. For example, a beer still can be used to process a fermented
10 product, which contains ethanol in an aqueous mixture, to produce an enriched ethanol-containing mixture that is then subjected to fractionation (e.g., fractional distillation or other like techniques). Next, the fractions containing the highest concentrations of ethanol can be passed through an adsorber to remove most, if not all, of the remaining water from the ethanol.

15

The following examples illustrate the invention:

EXAMPLES

20

METHODS

Molecular biology techniques and chemicals. Restriction enzymes and T4 DNA ligase were acquired from Fermentas. Antibiotics hygromycin (HG), phleomycin (phleo) and geneticin (G418) were acquired from Invivogen. pYL16 and nourseothricin
25 (nour) were acquired from Werner Bioagents. Ampicillin and kanamycin were acquired from Sigma-Aldrich.

For PCR amplifications, Phusion® High-Fidelity DNA Polymerase was used (Finnzymes). PCR fragments were sub-cloned using the TOPO® TA Cloning® Kit or the
30 Zero Blunt® TOPO® PCR Cloning Kit (both from Life Technologies). Oligonucleotides used for strain construction were purchased from Sigma-Aldrich.

Plasmids were amplified and maintained in chemically competent TOP10 cells (TOPO® TA Cloning® Kit, Life Technologies) following manufacturer's instructions.

Plasmids were isolated from *E. coli* mini cultures using the GeneJET™ Plasmid Miniprep Kit (Fermentas).

5

Genomic DNA was isolated from yeast using the YeaStar™ Genomic DNA Kit (ZymoResearch) following manufacturer's instructions.

Standard molecular biology and yeast genetics techniques were conducted according to textbooks including Sambrook *et al.* (1989) and Ausubel *et al.* (1995).

10

Strains and maintenance. For storage of the strains used in this work (Table 2), shake flask cultures were performed in rich medium (YP), consisting of 10 g l⁻¹ yeast extract (Oxoid) and 20 g l⁻¹ peptone (BD Difco), supplemented with either 2% glucose (YPD), 2% maltose (YPM), or 3% xylose (YPX). Cultures were maintained at 30 °C in an orbital shaker until cultures reached stationary growth phase. After adding glycerol to 30% (v/v), samples from shake-flask cultures were stored in 2 ml aliquots at -80 °C.

15

Table 2: Strains used or prepared herein

Strain	Genotype
RN1001	<i>Mat a, ura3-52, leu2-112, gre3::loxP, loxP-Ptpi:TAL1, loxP-Ptpi::RKI1, loxP-Ptpi-TKL1, loxP-Ptpi-RPE1, delta::Padh1XKSTTcyc1-LEU2, delta::URA3-Ptpi-xylA-Tcyc1</i>
RN1014	RN1001 + <i>in vivo</i> engineering on xylose and acetic acid
RN1041	RN1001 <i>his3::loxP</i>
RN1053	RN1041 <i>hxt2::loxP-kanMX-loxP, hxt367::loxP-hphMX-loxP, hxt145::loxP-natMX-loxP, gal2::loxP-zeoMX-loxP</i>
YD01227	RN1014 <i>glk1::lox72; hxx1::loxP; hxx2::lox72; gal1::loxP; his3::loxPnatMXloxP</i>

20

OD600 and HPLC analysis in shake flask culture. Shake flask cultures were sampled regularly during culture. For OD600 measurements, cultures were diluted appropriately for accurate measurement and optical density was measured at 600 nm wavelength in a Perkin Elmer Spectrophotometer λ2 instrument. Remaining sample was filtrated to separate medium from yeast.

25

The filtrate was inserted into the appropriate vials for HPLC analysis. The concentrations of glucose, xylose, glycerol, acetic acid and ethanol in the medium were determined using a Shimadzu HPLC system. The system is equipped with column oven

CTO-10A-vp and Autoinjector SIL-10AD-vp with a guard column (Bio-Rad H cartridge, Bio-Rad) and an Aminex HPX-87H column (300 x 7.8 mm; Bio-Rad). Elution took place at 80 °C with 5 mM H₂SO₄ at 0.6 mL/min. The eluate was monitored using a Refractive Index detector RID-10A (Shimadzu).

5 **Microwell plate culture for growth curve profiling.** For micro-well cultivation of strains, the Bioscreen C (Growth Curves Ltd.) was used. Overnight pre-cultures were pelleted, washed with demi water and diluted in demi water to twice the desired OD₆₀₀ for inoculation. Medium was prepared in twice the concentration as desired. In one well of a honeycomb wellplate, 150 µl medium was mixed with 150 µl cell suspension.
10 Measurements were conducted in triplicate. Settings for the Bioscreen C were maintained at 30°C incubation T, measurements every 15 min, shaking at type Continuous, amplitude Maximum, and speed Normal. Shaking was set to stop 5 sec before measurement.

Automated transformation and colony picking. For the generation of
15 transformation of a saturation mutagenesis library into the model strains shake-flask cultures were performed in either YPM for RN1053, or YPX for YD01227 (see below). Yeast cells were pelleted and, subsequently, used in an automated transformation protocol based on Schiestl and Gietz (1989). Transformation mixes were plated on selection medium consisting of yeast nitrogen base (Sigma-Aldrich; 6.7 g l⁻¹), agar (BD
20 Biosciences; 15 g l⁻¹), supplemented with either 2% maltose (RN1053 transformations) or 3% xylose (YD01227 transformations). Transformation plates were incubated at 30 °C, and after colony formation, colonies were re-plated using an automated process transferring colonies to 96 well microtiter plates (MTP) containing the above-referred selection media. MTPs with transformants were incubated at 30 °C until clear growth
25 was observed.

NMR analysis. For the quantification of glucose, xylose, glycerol, acetic acid and ethanol in the sample, 100 µl sample is transferred accurately into a suitable vial. Subsequently 100 µl internal standard solution, containing maleic acid (20 g/l), EDTA (40 g/l) and trace amounts of DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) in D₂O, and
30 450 µL D₂O is added.

 1D ¹H NMR spectra are recorded on a Bruker Avance III 700 MHz, equipped with a cryo-probe, using a pulse program with water suppression (power corresponding to 3 Hz) at a temperature of 27°C.

The analyte concentrations are calculated based on the following signals (δ relative to DSS):

- α -glucose peak at 5.22 ppm (d, 0.38 H, $J = 4$ Hz),
- α -xylose peak at 5.18 ppm (d, 0.37 H, $J = 4$ Hz),
- 5 • glycerol peak at 3.55 ppm (dd, 2 H, $J_{1,2} = 6$ Hz and $J_{1a,1b} = 12$ Hz)
- acetic acid peak at 1.91 ppm (s, 3 H)
- ethanol peak at 1.17 ppm (t, 3 H, $J = 7$ Hz)
- The signal user for the standard:
- Maleic acid peak at 6.05 ppm (s, 2H)

10

Example 1 - Hexose transporter gene deletions

Deletion cassettes construction. Primers used in plasmid constructions are shown in Table 3; generated plasmids are shown in Table 4. Schemes with restriction sites used for cloning and sites used to release deletion constructs from the plasmid backbone are shown in Table 5.

15

Table 3: Primers (oligonucleotides) used in the examples

SEQ ID NO:	Internal code	Primer	Sequence (5' → 3')	Gene(s)	Purpose
2	5034	Kanf	AAGCTTGCCCTCGTCCCCGCC	kanMX	Amplification kanMX
3	5035	Kanr	GTCGACACTGGATGGCGGCG	kanMX	Amplification kanMX
4	5116	lf2	ATTCTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTAAGCTTGCCCTCGTCCCCGCCG	loxP	Part of loxP flank
5	5118	lr2	CATACATTATACGAAGTTATGCGCGCTCTAGATATCGTCGACACTGGATGGCGGCG	loxP	Part of loxP flank
6	5115	lf1	ATCCGGACGTACGTATAACTTCGTATAGCATACATTATACGAAGTTATTCTAGTAACGGCCGCCA	loxP	Reamplification/full loxP flank
10	5117	lr1	TCATGACGTCTCGAGGCCTATAACTTCGTATAGCATACATTATACGAAGTTATGCGCGCT	loxP	Reamplification/full loxP flank
11	115	Natf	ACATGTAAAATGACCACTCTTGACGACACGGC	nat1	Amplification nat1
13	116	Natr	CAGTACTAGGGGCCAGGGCATGCTC	nat1	Amplification nat1

SEQ ID NO:	Internal code	Primer	Sequence (5'→3')	Gene(s)	Purpose
14	28	H3f	TGTACATCCGGAATTCTAGATTGGTGAGCGCTAGGAGTCACTGCC	HIS3	-
16	29	H3r	CTCGAGTATTTACACCCGCATATGATCCGTCG	HIS3	-
17	201	Hx2uf	GACTAGTACCGGTGTTTTCAAACCTAGCAACCCC	HXT2	Upstream flank
18	202	Hx2ur	CGTACGCGTCTTCCGGAAGGGTACCATCAGATTTCAATTGACC	HXT2	Upstream flank
19	203	Hx2df	GAAGACACTCGAGACGTCTTTGTCTGTGAAACCAAGGGC	HXT2	Downstream flank
20	204	Hx2dr	GTCGACGGGCCCTTATGTTGGTCTGTTTAGTATGGCCG	HXT2	Downstream flank
21	205	Hx3uf	AAGCGGCCGCACTAGTACCGGTGAAACAACCTAATAACGATGTGGGAC	HXT3	Upstream flank
22	206	Hx3ur	ATCCGACGTCTTCTCAAGAAACAGTTTGGGCGACG	HXT3	Upstream flank
23	210	Hx4df	AGAAGACGCTCGAGACGTCCCTTATGGGAAGAAGGTGTTTTGCC	HXT4	Downstream flank
24	211	Hx4dr	ATGGATCCTAGGGGTTCTTGACAGATAAAGTGGC	HXT4	Downstream flank
25	212	Hx5uf	AAGCGGCCGCACTAGTACATGTGAACCTGAAACGCTCATCAAGGC	HXT5	Upstream flank
26	213	Hx5ur	TTCGTACGCGTCTTCCGAGTAACATGAAACCAGTACCACG	HXT5	Upstream flank
27	229	Hx7df	AGAAGACCCCTCGAGACGTCCGACGCTGAAGAAATGACTCACG	HXT7	Downstream flank
28	230	Hx7dr	AGTCGACGGATCCGTAATTTTCTCTTTAAGTACGGGCG	HXT7	Downstream flank
29	243	Gal2ufn	AAGCGGCCGCACTAGTACCGGTGATCTATATTCGAAAGGGGCGG	GAL2	Upstream flank
30	244	Gal2urn	AACGTACGTCCGGATCATTAGAATCTTTGAGATTGTGCGCT	GAL2	Upstream flank
31	233	Ga2df	AGAAGACCCCTCGAGACGTCTTACCTTGAAATCTGAAGGCTGG	GAL2	Downstream flank
36	234	Ga2dr	GTGGATCCTAGGTAAACGGTACGAGAAAAGCTCCG	GAL2	Downstream flank
42	281	Hx3inr2	GCTCTTTTCACGGAGAAATTCGGG	HXT3-6-7	Integration check
43	289	Hx2inf	TCTTCGGGAAGTAGATAGGTGGC	HXT2	Integration check
38	290	Hx2inr	GAAGTAATCAGCCACAATACGCC	HXT2	Integration check
39	299	Hx4inr2	CCATACTATTGTGCGACTCAAGCGC	HXT5-1-4	Integration check
37	317	Hx5inf	GGGTAAATAGTTTTAGGGGCACGG	HXT5-1-4	Integration check
40	323	Hx7inr1	GATGAGAATCCTTGGAACCGC	HXT3-6-7	Integration check

SEQ ID NO:	Internal code	Primer	Sequence (5'→3')	Gene(s)	Purpose
41	324	Ga2in1	TCAATTCGGAAAGCTTCCTCCGG	GAL2	Integration check
44	325	Ga2in1	CAGTGATAGTTTGGTTCGAGCGG	GAL2	Integration check
45	838	Glk1-psuc227f	ATGTCATTCGACGACTTACACAAAGCCACTGAGAGAGCGTCATCCAGGCC GTCGACCTCGAGTACCGTTG	GLK1	Hexokinase flank/Bipartite cassette
46	834	Hxk2-psuc227f	GCCAGAAAGGGTCCATGGCCGATGTGCCAAAGGAATTGATGCAACAAATCC GTCGACCTCGAGTACCGTTG	HXK2	Hexokinase flank/Bipartite cassette
47	645	pSUC227r	GCAATTCGGCTATACGTAAC		Bipartite cassette
48	839	Glk1-psuc225r	CAATCTCAAGTGACCTTCCTCTCACCTCGGCACCCAGGGTGACAAGCC GGATCCTACCGTTCTGATAGC	GLK1	Hexokinase flank/Bipartite cassette
49	835	Hxk2-psuc225r	GCCAGAAAGGGTCCATGGCCGATGTGCCAAAGGAATTGATGCAACAAATCC GTCGACCTCGAGTACCGTTG	HXK2	Hexokinase flank/Bipartite cassette
50	646	pSUC225f	CGTTCACCTCATGGAAAATAGC		Bipartite cassette
51	846	Hxk1_loxP_f	ATGGTTCATTTAGGTCCAAAGAAACACAGGCTAGAAAGGGTCCATGGCCG GATCCACTAGCATAACTTCG	HXK1	Hexokinase flank /DRM cassette
52	847	Hxk1_loxP_r	ATGGTTCATTTAGGTCCAAAGAAACACAGGCTAGAAAGGGTCCATGGCCG GATCCACTAGCATAACTTCG	HXK1	Hexokinase flank /DRM cassette
53	848	Gal1_loxP_f	ATGACTAAATCTCATTCAGAAGAAGTGATTGTACCTGAGTTCAATTCTAGCGG ATCCACTAGCATAACTTCG	GAL1	Hexokinase flank /DRM cassette
54	849	Gal1_loxP_r	TTATAATTCATATAGACAGCTGCCAATGCTGGTTTAGAGACGATGATGTTG GGCCGCCAGTGATG	GAL1	Hexokinase flank /DRM cassette

5 Table 4: Plasmids used in the strain construction

Number	Construct	Purpose	SEQ ID NO:
pRN201	pCR-BLUNT-loxP-kanMX-loxP	Dominant resistance marker	8
pRN251	pCR-BLUNT-loxP-hphMX-loxP	Dominant resistance marker	9
pRN365	pCR-BLUNT-loxP-naIMX-loxP	Dominant resistance marker	10
pRN447	pCR-BLUNT-loxP-zeoMX-loxP	Dominant resistance marker	14
pRN247	pCR-BLUNT-his3:loxP-kanMX-loxP	HIS3 deletion construct	16
pRN485	pCR-BLUNT-gal2:loxP-zeoMX-loxP	GAL2 deletion construct	33

pRN566	pCR-BLUNT- <i>hxt367:loxP-hphMX-loxP</i>	<i>HXT3-HXT6-HXT7</i> cluster deletion construct	34
pRN569	pCR-BLUNT- <i>hxt514:loxP-naMX-loxP</i>	<i>HXT5-HXT1-HXT4</i> cluster deletion construct	35
pRN635	pCR-BLUNT- <i>hxt2:loxP-kanMX-loxP</i>	<i>HXT2</i> deletion construct	36
pRN993	pRN978-P _{HXT7(-491)} - <i>GAL2-T_{ADHI}</i>	<i>GAL2</i> expression vector	58
pDB1250	pRN978-P _{HXT7(-491)} -synt.wt- <i>GAL2-T_{ADHI}</i>	Synthetic wildtype <i>GAL2</i> expression vector	59
pRN187	pCRE-zeoMX (based on pSH65)	CRE recombinase expression vector	60
pRN486	pCR-BLUNT-HIS3::loxPnatMXloxP	HIS3 deletion construct	61

Table 5: Cloning scheme

Construct	Fragment	Cloning Sites	Release Sites
pRN247	<i>HIS3</i> upstream	<i>SacI-DraI</i>	<i>XhoI-BsrGI</i>
	<i>loxP-kanMX-loxP</i>	<i>StuI-BsiWI</i>	
	<i>HIS3</i> downstream	<i>BsiWI-ApaI</i>	
pRN485	<i>GAL2</i> upstream	<i>NotI-BsiWI</i>	<i>BamHI-SpeI-PmlI</i>
	<i>loxP-zeoMX-loxP</i>	<i>BsiWI-XhoI</i>	
	<i>GAL2</i> downstream	<i>XhoI-BamHI</i>	
pRN566	<i>HXT3</i> upstream	<i>SpeI-BsiWI</i>	<i>BamHI-AgeI-NaeI</i>
	<i>loxP-hphMX-loxP</i>	<i>BsiWI-XhoI</i>	
	<i>HXT7</i> downstream	<i>XhoI-BamHI</i>	
pRN569	<i>HXT5</i> upstream	<i>NotI-BspEI</i>	<i>BamHI-NotI-ApaLI</i>
	<i>loxP-naMX-loxP</i>	<i>BspEI-XhoI</i>	
	<i>HXT4</i> downstream	<i>XhoI-BamHI</i>	
pRN635	<i>HXT2</i> upstream	<i>SpeI-BsiWI</i>	<i>AgeI-NotI-BspHI</i>
	<i>loxP-kanMX-loxP</i>	<i>BsiWI-XhoI</i>	
	<i>HXT2</i> downstream	<i>XhoI-EcoRI</i>	

5

The *kanMX* marker was amplified from the plasmid pFA6-kanMX4 (http://www-sequence.stanford.edu/group/yeast_deletion_project/kanmx4.txt) using primers SEQ ID NO's 1 and 2. Subsequently, the *kanMX* marker was floxed through adding *loxP* flanks by PCR amplification with primers SEQ ID NO's 3 and 4. Re-amplification was done with

primers SEQ ID NO's 5 and 6. The resulting *loxP-kanMX-loxP* fragment was cloned in pCR-BLUNT resulting in pRN201 (SEQ ID NO: 7).

For the construction of pRN251 (SEQ ID NO: 8), hphMX was isolated from pGRE3:hphMX (Kuyper *et al*, 2005). To delete a *MluI* site as appropriate restriction site in the vicinity of hphMX, pGRE3:hphMX was cut with *Eco32I* and re-ligated. Subsequently, hphMX was cloned as *XhoI-MluI* fragment into pRN201 digested with *SaII* and *MluI* to replace kanMX.

For the construction of pRN365 (SEQ ID NO: 9), the *Streptomyces noursei nat1* gene was PCR-amplified from pYL16 (Werner Bioagents) using primers with SEQ ID NO's 10 and -11. The *PstI-ScaI nat1* fragment together with the *Acc65I-NcoI* pRN201-fragment were cloned into pRN201, already linearized with *Acc65I* and *ScaI*, in order to replace *kanR* for *nat1*.

For the construction of pRN447 (SEQ ID NO: 12), pRN201 was digested with *PmlI*. This served two ends. Firstly, the *Streptoalloteichus hindustanus ble* (zeocin or phleomycin resistance gene) ORF was isolated, and secondly, after re-ligation of the *PmlI*-digested pRN201 an *NcoI* site was deleted. Subsequently, *ble* as *NcoI-PmlI* fragment and part of pRN201 as *BamHI-NcoI* vector fragment were cloned into the re-ligated pRN201 (missing *ble*), digested with *BamHI* and *ScaI* resulting in pRN447.

For the *HIS3* deletion construct, primers SEQ ID13 and-14 were used to amplify the *HIS3* locus from yeast genomic DNA. Sites used to cut out the *HIS3* flanks and to ligate these to the floxed kanMX marker are shown in Table 5. The ligation product was digested with *SacI* and *ApaI* and cloned into pCR-BLUNT digested with *SacI* and *ApaI*. The resulting plasmid is pRN247 (SEQ ID NO: 15).

For the deletion of the eight main hexose transporters (*HXT1-7* and *GAL2*) in *S. cerevisiae*, four deletion constructs were generated (see Table 4). Each deletion construct contained a different floxed dominant resistance marker. For each *HXT* gene 400-700 bp flanks were amplified using the primers listed in Table 3 (SEQ ID NO's 16-31) using RN1001 genomic DNA as template. The upstream flank, the dominant resistance marker and the downstream marker were ligated using the fragments and cloning sites listed under Table 5. The ligations were amplified using the forward primers of the upstream flank and the reverse primer of the downstream flank (primer combinations SEQ ID NO's 16+19, SEQ ID NO's 20+27, SEQ ID NO's 24+23, and SEQ ID NO's 28+31). The fused PCR fragments were cloned into pCR-BLUNT to obtain pRN485, pRN566, pRN569, pRN635 (SEQ ID NO's 32-35, respectively). To

obtain high yields of plasmid DNA, the plasmids were isolated from 50 mL *E. coli* cultures using NucleoBond® Xtra Midi kit (Bioké, Leiden, the Netherlands). Before transformation to yeast, deletion constructs were released from plasmid backbone by digestion with the release restriction sites listed in Table 5.

5

Strain construction. The xylose-fermenting strain RN1001 was made histidine auxotroph by the insertion of *loxP-kanMX-loxP* (released from pRN247; SEQ NO ID15) at the *HIS3* locus. Subsequently, the marker was removed through transient expression of plasmid pRN187 (derived from pSH65 expressing galactose-inducible *cre* recombinase; SEQ ID NO 60). Introduction of pRN187 was selected on phleo and *CRE* recombinase expression was induced on YP-medium supplemented with galactose. The resulting *his3:loxP* strain was named RN1041. The hexose transporters were deleted in the following order: 1) *HXT3-HXT6-HXT7* cluster, 2) *HXT5-HXT1-HXT4* cluster, 3) *GAL2*, 4) *HXT2*. The deletion constructs were linearized or released from the plasmid backbone by cutting with the enzyme combinations listed in Table 5 and these were integrated in the genome of RN1041. All transformations were plated on yeast extract (10 g/L), peptone (20g/L) agar (15g/L) medium supplemented with 20g/L maltose. Maltose was added to the medium, because the uptake of this disaccharide goes via an alternative transport system than the glucose transport system (Wieczorke *et al*, 1999). With each deletion of a (cluster of) *HXT* gene(s), an additional marker was inserted in the order: 1) hphMX, 2) natMX, 3) zeoMX, 4) kanMX. With each inserted additional marker the respective antibiotic was additionally supplemented to the medium in the following order: 1) HG, 2) HG and nour, 3) HG, nour and phleo, 4) HG, nour, phleo and G418. After integration of all four deletion constructs, a single colony was isolated under selection of all four antibiotics. Correct integrations were verified by PCR analysis on genomic DNA isolates. Primers outside of the integration site were used (combinations SEQ ID NO:'s 36+37, SEQ ID NO:'s 38+39, SEQ ID NO:'s 40+41, SEQ ID NO:'s 42+43; sequences listed in Table 3).

30

Strain characterization. To characterize the (intermediate) hexose transporter strains, shake flask cultures were performed. Cultures were inoculated at OD600=0.1. The resulting strain, RN1053 (Δ hxt1-7;gal2-mutant RN1041; see Table 2 for exact genotype), showed a retarded growth pattern on Verduyn-urea (mineral medium according to Verduyn using urea as nitrogen source; Luttik *et al*, 2001) supplemented

with 0.2 g l⁻¹ histidine (Sigma-Aldrich; Verduyn-urea-his; to complement for the histidine auxotrophy) and 15 g l⁻¹ glucose and 20 g l⁻¹ xylose only starting to grow slowly on glucose only after 60 hours; interestingly, when glucose was present in the medium xylose was finished as well after 150 hrs (Figure 1) indicating that one or more of the cryptic hexose transporter genes (*HXT8-17*) was induced on glucose and facilitated xylose transport (Figure 1). However, on xylose as sole carbon source RN1053 did not grow on Verduyn-urea (+ 20 g l⁻¹ % xylose during the culturing period (Figure 2) indicating the strain is useable as model strain for testing putative xylose transporters. RN10153 was further maintained on YPM.

Example 2 – Hexokinase gene deletions

Deletion cassettes construction. For deletion of hexokinase genes oligonucleotides were designed (SEQ ID NO:s in Table 3) comprised of 60 nucleotide flanking sequences homologous to the hexokinase gene locus and of 20 nucleotides homologous to a floxed dominant resistance marker cassette. The oligonucleotides were used to amplify the deletion constructs. Subsequent PCR products were column filter-purified (Fermentas GeneJet Kit) and used for transformations experiments. Three types of deletion cassettes were used: firstly, for *GLK1* and *HXK2* deletions a bipartite system was used. One fragment consisted of a lox66 site, KanMX, *GAL1* promoter upstream of *CRE*, and the 5'-part of *CRE* (*CRE1*) amplified from pSUC227 with one gene-specific primer (SEQ ID NO: 44 for *GLK1* and SEQ ID NO: 45 for *HXK2*) and one pSUC227-specific primer (SEQ ID NO: 46); the second fragment consisted of the 3'-part of *CRE* (*CRE2*) with overlap on *CRE1*, and a lox71 site, amplified from pSUC225 with again one gene-specific primers (SEQ ID NO: 47 for *GLK1* and SEQ ID48 for *HXK2*) and one pSUC225-specific primer SEQ ID NO: 49. Through homologous recombination the two fragments integrate as lox66-kanMX-CRE-lox71 at the hexokinase locus (pSUC225 and pSUC227 sequences and method provided in PCT/EP2013/055047). Secondly, for *HXK1* (primers SEQ ID NO:s 50-51) and *GAL1* (primers SEQ ID NO:s 52-53) deletions, a floxed dominant resistance marker (DRM) was amplified with flanking sequences homologous to the respective hexokinase to replace the coding region at the locus; as templates for the PCR amplifications of the DRM cassettes pRN774 (loxP-hphMX-loxP; SEQ ID NO: 54) and pRN775 (loxP-natMx-loxP; SEQ ID NO: 55) were used, respectively. Thirdly, for *HIS3* deletion to allow for complementation of the auxotrophic

phenotype by transporter episomal plasmids, a similar construct with *HIS3*-homologous flanks was integrated as was used to generate RN1041 (RN1001-*his3*:loxP, in strain family RN1053; see above Example 1). In this case the construct beared natMX as dominant resistance marker instead of kanMX (SEQ ID NO 61).

5 **Strain construction.** For the generation of a strain incapable of hexose metabolism but capable of hexose transport, hexokinase gene deletions were made in the xylose-fermenting strain RN1014 (Table 2; Figure 3 for deletion scheme).

10 As mentioned, in the case of *GLK1* and *HXK2*, the disruption cassettes were bipartite. Through homologous recombination the two fragments integrate as lox66-kanMX-CRE-lox71 at the hexokinase locus. The integration was selected on YPD supplemented with G418. The disruption cassettes for *HXK1* and *GAL1* consisted of one fragment: either loxP-natMX-loxP or loxP-hphMX-loxP, respectively. RN1014 was transformed with the purified PCR products and the integration was selected on the appropriate antibiotic. Additionally, *HIS3* was disrupted with a similar construct (SEQ ID NO: 61) used for the generation of RN1053. In this case natMX was the selection marker instead of kanMX.

15 The genes were deleted in the following order: 1) *GLK1*, 2) *HXK1*, 3) *GAL1*, 4) *HXK2* and 5) *HIS3*. After the deletion of *GLK1* and *HXK1* both markers were recycled by galactose-induced Cre-mediated recombination. After deletion of *HXK2* the intermediate strain was maintained on xylose-containing rich medium (YPX). After *HIS3* deletion the integrated hphMX and kanMX markers were removed by galactose-induced CRE recombination. To ensure growth of the strain, 2% xylose was added to YP 2% galactose + nourseothricin (YPGX). Selection on nourseothricin ensured maintenance of the natMX marker at the *HIS3* locus leaving a selection trait to be used possibly later on. After single colony isolation, the strain was verified for its deletions and delta sequence profile by colony PCR, and named YD01227.

20 **Strain characterization.** In aerobic shake flask culture experiments using Verduyn-urea-his, another colony with the same quadruple hexokinase knockout (KO) genotype (col 2) as YD01227 was characterized in a pre-screen for its ability to consume xylose in the absence and presence an excess of glucose (10%), and for its ability to consume glucose (Figure 4). Cultures were inoculated at OD₆₀₀=0.1. Both RN1014 and the quadruple hexokinase KO are able to grow on and consume xylose (data not shown). And expectedly, the quadruple hexokinase KO does neither grow on nor

30

consumes glucose, whereas RN1014 does (data not shown). Furthermore, the excess of glucose (10%) prevents the growth on and consumption of xylose for at least 96 hours, whereas RN1014 utilizes xylose (Figure 4).

In Bioscreen C experiments, YD01227 and the aforementioned col 2 were
5 screened for growth on Verduyn-urea-his supplemented with different sugars and sugar mixtures. Cultures were inoculated at OD600=0.05. YD01227 grew on xylose but was not able to grow on glucose, maltose or galactose (data not shown). The glucose-xylose mixtures were screened to support a choice of medium composition suitable for the screen for pentose-specific transporters. As seen in Figure 5, with a ratio of glucose:
10 xylose of 5:1 showed the optimal inhibition of growth on xylose for strain YD01227. This was the case for both a high (10:2), as for a low sugar load (2.5: 0.5). YD01227 was further maintained on YPX for storage and handling.

Example 3. GAL2 saturation mutagenesis library and other constructs.

15 A synthetic DNA construct for wild type (WT) GAL2 was ordered at GeneArt (SEQ ID NO: 56; Invitrogen) and was used as template for site-directed mutagenesis. The synthetic WT GAL2 DNA construct was cloned into pRN993 (SEQ ID NO: 57) as *Xba*I-*Bss*HI fragment exchanging another ORF for the synthetic WT GAL2 construct to generate pDB1250 (SEQ ID NO: 58). pDB1250 is a yeast shuttle vector based on
20 pRS313, bearing as most prominent features besides the synthetically made GAL2 ORF: 1) a *HIS3* expression cassette to complement the histidine auxotrophy, 2) a CEN.ARSH to maintain 1-2 copies (low copy number) of the expression vector in yeast cells, 3) truncated *HXT7* promoter (-491 bp) resulting in medium expression levels of downstream ORF, 4) *ADH1* terminator, and 5) ampicillin resistance gene (*amp^r*) for
25 selection in *E. coli* TOP10 cells (see above) for cloning purposes (Figure 6). The saturation mutagenesis library for at least 13 non-wildtype amino acid changes on 30 amino acid positions in Gal2p was ordered from Invitrogen Life Technologies. (<http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Cloning/gene-synthesis/directed-evolution/GeneArt-Site-Saturation-Mutagenesis.html>); for positions and amino acid changes to the wildtype
30 Gal2p amino acid sequence (SEQ ID NO: 59) see Table 6.

Table 6: GAL2 Single Site Saturation Mutagenesis Library

# sites	Gal2		# non wt AA	Non wt AA
	Position	WT AA		
1	85	F	15	A, C, D, E, G, H, K, M, N, P, Q, R, S, T, V
2	89	T	18	A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, V, W, Y
3	187	V	17	A, C, D, E, F, G, H, I, K, L, M, Q, R, S, T, W, Y
4	191	A	19	C, D, E, F, G, H, K, L, M, N, P, Q, R, S, T, V, W, Y
5	214	Y	15	A, C, G, I, K, L, M, N, P, Q, R, S, T, V, W
6	215	Q	16	A, C, D, E, F, G, H, I, L, M, N, R, S, V, W, Y
7	218	I	15	A, C, D, E, G, H, K, L, M, N, R, S, T, V, W
8	219	T	15	A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, V, W
9	222	I	15	A, C, D, E, G, H, K, M, P, R, S, T, V, W, Y
10	226	Y	19	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W
11	338	Q	18	A, C, D, E, F, G, H, I, L, M, N, P, R, S, T, V, W, Y
12	339	M	19	A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, Y
13	341	Q	13	D, E, H, I, K, L, M, N, R, S, T, V, Y
14	342	Q	19	A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, Y
15	343	L	19	A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, Y
16	346	N	19	A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W, Y
17	347	N	13	A, D, E, G, H, I, K, L, Q, R, S, T, V
18	350	F	14	D, E, H, I, K, L, M, N, Q, R, S, T, V, Y
19	373	G	19	A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y
20	376	N	18	A, C, D, E, F, H, I, K, L, M, P, Q, R, S, T, V, W, Y
21	380	T	18	A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, V, W, Y
22	383	S	17	A, C, D, E, F, G, I, K, L, M, N, Q, R, T, V, W, Y
23	444	F	19	A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y
24	446	Y	13	D, E, H, I, K, L, M, N, Q, R, S, T, V
25	448	T	18	A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, V, W, Y
26	449	T	19	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, Y
27	451	A	18	C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y
28	455	W	18	A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, Y
29	478	N	16	A, C, D, F, G, H, I, K, L, P, Q, R, V, W, Y
30	479	W	18	C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, Y

- 5 Resulting constructs were inserted by custom cloning at GeneArt (Invitrogen, Regensburg, Germany) in pDB1250 (SEQ ID NO: ID58).

Example 4: Glucose Transport Activity Counter-screening

- Aim.** Using the Glucose Transport Activity Counter (GTAC)-screen, i.e. transforming hexokinase-mutant strain YD01227 as host to introduce GAL2 variants and screening the resulting transformants on medium for growth on and consumption of xylose in the presence of a 5 times higher amount of glucose, mutations can be identified that favour xylose above in the presence of a surplus glucose in the Gal2p variant (i.e. higher affinity for xylose than for glucose reduction or more preferably full removal of glucose transport capability, while keeping xylose transport capability more or less intact).

Transformation and colony picking. YD01227 was transformed with a total of 497 constructs, each one bearing a *GAL2* mutant. One construct with wild-type *GAL2*

sequence (pDB1250; SEQ ID NO: 58) was included as control. For each transformation, 3 colonies, when available, were re-plated to agar medium MTPs amounting to 1450 transformants. The 1450 transformants were screened in three parts. For each part of the screening, wild-type *GAL2* was included as control.

Pre-culture and screen. For pre-culture, transformants were transferred by automated process from selection agar medium MTPs to 96 well half-deepwell plates (96HDWP) containing liquid selection medium consisting of mineral medium (according to Verduyn, with urea as nitrogen source) supplemented with 3% xylose. The 96 HDWPs were cultured for 3 days of pre-culture in an orbital shaker at 30 °C and 750 rpm. Subsequently, for each well 20 µl of culture was transferred by automated process to 24 deepwell plates (24 DWPs) containing 2.5 ml Verduyn-urea supplemented with the sugar mixture glucose:xylose in a ratio 5:1 with the following concentrations: 10 g l⁻¹ glucose and 2 g l⁻¹ xylose (YD10-medium). For each of the three sampling points a series of 24DWPs was inoculated. On each 24DWP, one RN1001 growth control was inoculated to have an indication of plate variation effects. After 24 hours, 72 hours and 96 hours automatic sampling and transfer to 96HDWP was conducted for automated OD-measurement at 600 nm wavelength.

After the last OD measurement after 96 hours, cells were pelleted after centrifugation and 100 µl supernatant was collected by automated process for flow-NMR analysis residual constituents in the medium after culture (see above for method description). For each construct and each time point, the measured residual glucose and xylose concentrations for the different replicates were averaged. In order to compare the different parts of the screen, a relative value was calculated based on the difference to the wildtype residual xylose concentration measured in the particular part of the GTAC screen, according to the following formula:

$$\text{RelXyl} = \frac{(\text{residual xylose}_{\text{Gal2-wildtype}} - \text{residual xylose}_{\text{Gal2-mutant}})}{\text{residual xylose}_{\text{Gal2-wildtype}}} \times 100\%$$

Results. Whereas RN1001 displayed complete consumption of both glucose and xylose, YD01227 transformants of wildtype and mutagenesis library *GAL2* constructs did not consume glucose and displayed a spectrum of residual xylose concentrations representing their individual ability to consume xylose in the presence of glucose (Figure 7A). As shown in Figure 7B, the xylose consumption displayed a high correlation

($R^2=0.93$) with the growth measurements (OD600). Since the screen was conducted in aerobic conditions, and little ethanol formation was shown (data not shown) and growth displayed high correlation with the xylose consumption, the residual xylose concentrations were used as main parameter to compare GAL2 variants with wildtype.

5 The comparison of all tested mutant *GAL2* variants versus wildtype *GAL2* on xylose consumption (average RelXyl) at 96 hours is listed in Table 7. All mutations affecting glucose transport to the benefit of xylose transport have RelXyl>0 (Table 7). The TOP positions to target for a second round of mutagenesis were sorted based on their average RelXyl score per mutation; specific mutations were sorted on preference based

10 on RelXyl score as well (Table 8).

Table 7. Xylose consumption (RelXyl) of GTAC Screen. Wildtype GAL2 construct pDB1250 was set to 0. RelOD600 values are the averages of 1-3 replicates and are relative values compared to wildtype.

15

Pos	wtAA	MutAA	RelXyl
85	F	A	1,45
		C	0,66
		D	-1,89
		E	-13,09
		G	-1,06
		H	1,76
		K	-20,81
		M	7,03
		N	1,96
		P	2,83
		Q	37,68
		R	-1,93
		S	19,81
		T	50,51
		V	32,76
89	T	A	11,68
		C	13,22
		D	19,68
		F	-9,97
		H	25,59
		I	-17,32
		K	-24,84

Pos	wtAA	MutAA	RelXyl
		L	12,38
		M	-8,38
		N	9,88
		Q	-16,03
		R	-16,11
		S	-11,44
		V	41,06
		Y	-2,89
187	V	A	26,84
		C	9,98
		D	3,76
		E	6,21
		F	14,53
		G	10,01
		H	8,32
		I	6,69
		K	4,09
		L	4,12
		M	1,85
		Q	10,17
		R	3,05
		S	7,54
		T	10,41
		W	1,81
		Y	1,22
191	A	C	4,70
		D	12,43
		E	-2,61
		F	5,49
		G	15,83
		H	2,83
		I	11,93
		K	1,94
		L	-8,25
		M	-14,37
		N	20,35
		P	14,89
		Q	2,94
		R	-1,09
		S	3,09
		T	10,61

Pos	wtAA	MutAA	RelXyl
		V	4,52
		W	2,49
		Y	1,93
214	Y	A	-12,81
		C	-5,85
		G	18,41
		I	-12,07
		K	-4,66
		L	35,33
		M	23,92
		N	32,71
		P	18,44
		Q	39,94
		R	25,50
		S	14,91
		V	55,48
		W	47,35
215	Q	A	2,01
		C	9,90
		D	13,62
		F	10,12
		G	-9,62
		H	-18,18
		I	17,72
		L	24,63
		M	21,59
		N	-11,47
		S	2,14
		V	-8,22
		W	-20,47
		Y	-8,07
218	I	A	17,98
		C	5,52
		D	4,64
		E	8,44
		G	4,62
		H	9,88
		K	27,86
		L	9,73
		M	5,80
		N	24,31

Pos	wtAA	MutAA	RelXyl
		R	-0,73
		S	24,96
		T	-2,65
		V	7,69
		W	-2,10
219	T	A	52,55
		C	25,05
		D	18,04
		E	-14,58
		F	43,86
		G	46,45
		I	-11,07
		K	-4,44
		M	-18,28
		N	7,85
		Q	-22,81
		R	-24,62
		S	4,90
		W	-16,77
222	I	A	11,37
		C	13,45
		D	20,91
		E	3,35
		G	3,75
		H	9,73
		K	-4,92
		M	4,39
		P	0,97
		S	1,85
		T	-1,59
		V	8,42
		W	-8,04
		Y	7,87
226	Y	A	63,63
		C	62,07
		D	81,68
		E	70,43
		G	39,06
		H	5,10
		I	32,79
		K	-1,86

Pos	wtAA	MutAA	RelXyl
		L	42,48
		M	74,04
		N	62,81
		P	49,78
		Q	56,23
		R	65,53
		S	14,97
		T	7,27
		V	-11,15
		W	44,48
338	Q	A	38,18
		C	52,66
		D	33,17
		E	46,93
		F	46,65
		G	13,80
		H	26,28
		I	28,32
		L	33,08
		M	-2,53
		N	22,26
		P	3,96
		R	39,83
		S	3,96
		T	11,84
		V	24,84
		W	35,55
		Y	19,90
339	M	A	50,56
		C	30,36
		D	46,71
		E	10,48
		F	63,75
		G	73,03
		H	67,99
		I	40,96
		K	71,92
		L	67,43
		N	83,73
		P	12,77
		Q	77,74

Pos	wtAA	MutAA	RelXyl
		R	70,25
		S	81,65
		T	57,02
		V	82,35
		W	-23,02
		Y	37,43
341	Q	D	-20,08
		E	-3,38
		H	-22,19
		L	46,71
		M	23,17
		N	11,77
		R	-5,18
		S	15,78
		T	28,67
		Y	19,78
342	Q	A	61,47
		C	64,63
		D	48,62
		E	34,90
		F	45,61
		G	14,10
		H	46,28
		I	16,81
		K	41,75
		L	37,38
		M	27,70
		N	32,30
		P	44,88
		R	23,95
		S	68,48
		T	50,54
		V	24,39
		W	59,43
		Y	76,45
343	L	A	9,67
		C	8,74
		D	9,41
		E	2,39
		F	15,49
		G	11,87

Pos	wtAA	MutAA	RelXyl
		H	16,96
		I	10,67
		K	40,15
		M	16,94
		N	14,03
		P	26,08
		Q	8,73
		R	29,11
		S	20,28
		T	14,38
		V	-6,06
		W	7,76
		Y	20,43
346	N	A	21,99
		C	17,11
		D	8,88
		E	-8,32
		F	9,91
		G	8,42
		H	26,01
		I	17,12
		K	17,48
		L	6,48
		M	13,19
		Q	19,93
		R	7,94
		S	23,37
		T	23,19
		V	47,09
		W	45,58
		Y	41,70
347	N	A	44,64
		D	26,67
		E	25,36
		G	-22,71
		H	11,18
		I	27,33
		K	-22,05
		L	7,79
		Q	-14,96
		R	-15,36

Pos	wtAA	MutAA	RelXyl
		S	-15,56
		T	-19,33
		V	-28,12
350	Y	D	2,93
		E	-13,15
		H	-19,99
		I	-10,76
		K	3,00
		L	40,49
		M	5,01
		N	13,57
		Q	0,66
		R	57,58
		S	15,25
		T	21,64
		V	13,39
		Y	-7,36
373	G	A	25,90
		C	36,56
		D	36,22
		E	41,01
		F	15,15
		H	6,96
		I	9,10
		K	20,72
		L	51,94
		M	35,03
		N	38,49
		P	35,75
		Q	5,50
		R	8,35
		S	11,62
		T	10,67
		V	21,91
		W	22,96
		Y	31,04
376	N	A	36,70
		C	78,63
		D	19,23
		E	20,15
		F	92,62

Pos	wtAA	MutAA	RelXyl
		H	39,30
		I	99,46
		K	7,53
		L	93,69
		M	97,33
		P	51,70
		Q	24,24
		R	27,86
		S	59,39
		T	99,01
		V	96,57
		W	13,51
		Y	43,95
380	T	A	53,06
		C	42,34
		D	37,28
		E	-29,92
		F	54,67
		G	34,22
		I	47,40
		K	42,12
		L	39,25
		M	74,78
		N	49,79
		P	43,28
		Q	43,31
		R	42,31
		S	42,03
		V	42,53
		W	51,16
		Y	25,09
383	S	A	57,21
		C	67,78
		D	29,42
		E	54,55
		F	52,71
		G	39,71
		I	25,88
		K	49,19
		L	37,72
		M	52,64

Pos	wtAA	MutAA	RelXyl
		N	58,14
		Q	83,94
		R	61,78
		T	71,85
		V	19,38
		W	65,32
		Y	62,76
444	F	A	32,41
		C	14,64
		D	13,21
		E	16,38
		G	12,78
		H	25,62
		I	31,95
		K	28,79
		L	39,43
		M	17,30
		N	5,43
		P	15,65
		Q	19,63
		R	0,06
		S	10,94
		T	7,44
		V	37,29
		W	2,94
		Y	32,56
446	F	D	10,80
		E	10,63
		H	16,93
		I	8,65
		K	-3,40
		L	22,78
		M	18,62
		N	14,99
		Q	34,25
		R	-8,44
		S	-20,59
		T	-16,07
		V	-9,48
448	T	A	76,95
		C	58,53

Pos	wtAA	MutAA	RelXyl
		D	66,10
		E	53,25
		F	61,94
		G	65,05
		H	67,00
		I	58,01
		K	54,82
		L	63,49
		M	77,64
		P	64,39
		Q	58,15
		R	33,80
		S	59,40
		V	53,50
		W	56,69
		Y	47,73
449	T	A	57,64
		C	50,44
		D	55,89
		E	37,14
		F	92,65
		G	61,00
		H	82,47
		I	71,71
		K	79,30
		L	81,03
		M	89,29
		N	84,60
		P	37,06
		Q	81,19
		R	82,14
		S	83,15
		V	74,98
		W	80,75
		Y	79,80
451	T	C	10,57
		D	9,68
		E	15,39
		F	13,38
		G	43,34
		H	7,65

Pos	wtAA	MutAA	RelXyl
		I	17,96
		K	8,59
		L	11,00
		M	5,68
		N	14,56
		P	11,89
		Q	8,63
		R	4,46
		S	13,08
		T	2,10
		V	4,99
		W	15,25
		Y	1,62
455	W	A	1,31
		C	1,19
		D	-6,13
		E	19,68
		F	-1,57
		G	0,73
		I	-2,70
		K	6,91
		L	19,08
		M	19,92
		N	19,10
		P	14,76
		Q	3,03
		R	-14,86
		S	8,05
		T	19,75
		V	10,66
		Y	7,30
478	N	A	-11,71
		C	-1,14
		D	-3,08
		F	30,35
		G	8,29
		H	-15,02
		I	10,06
		K	9,95
		L	-3,27
		P	30,98

Pos	wtAA	MutAA	RelXyl
		Q	4,37
		R	11,41
		V	21,91
		W	3,25
		Y	4,98
479	W	C	13,20
		D	6,12
		E	11,20
		F	1,41
		G	1,34
		H	6,90
		I	1,50
		K	5,54
		L	7,80
		M	5,47
		N	2,20
		P	7,07
		Q	15,50
		R	1,63
		S	18,50
		T	21,51
		V	5,89
		Y	8,63

Table 8. Gal2p positions and mutations identified in the GTAC screen are ordered on preference. SCORE TOP HIT is the RelXyl score for the amino acid substitution within a position yielding the clearest improvement compared to wildtype Gal2p This allows a sorting of the most influential mutations and relevant positions to target in Gal2p to eliminate glucose affinity.

	GTAC	Mutations					
	Position	wt AA	Most preferred	Preferred	Least preferred	Inactive	SCORE TOP HIT
Most preferred	376	N	FILMTV	ACHPSY	DEKQRW		99,46
	449	T	FM	HIKLNQRSVWY	ACDEGP		92,65
	383	S	CQRTWY	AEFKMN	DGILV		83,94
	339	M	NQSV	AFGHKLRT	CDEIPY	W	83,73
	226	Y	ACDMNR	EQ	GHILPSTW	KV	81,68
	448	T	AM	CDEFGHIKLPQSVW	RY		77,64

	GTAC	Mutations					
	342	Q	ACSWY	DFHKPT	EGILMNRV		76,45
	380	T	AFMW	CDGIKLNQRSV	Y	E	74,78
	350	F	LR	NSTV	DKMQ	EHIY	57,58
	214	Y	VW	LNQ	GMPRS	ACIK	55,48
	338	Q	CEF	ADHILRW	GNPSTVY	M	52,66
	219	T	AFG	CD	NS	EIKMQRW	52,55
	373	G	L	ACDEMNPVWY	FHIKQRST		51,94
	85	F	QTV	S	ACHMNP	DEGKR	50,51
Preferred	346	N	VWY	AHST	CDFGIKLMQR	E	47,09
	341	Q	L	MT	NSY	DEHR	46,71
	347	N	A	DEI	HL	GKQRSTV	44,64
	451	A	G	CDEFILNPSW	HKMQRTVY		43,34
	89	T	DHV	ACL	N	FIKMQRSY	41,06
	343	L	K	PRSY	ACDEFGHIMNQTW	V	40,15
	444	F	AILVY	HK	CDEGMNPQSTW	R	39,43
	446	F	Q	HLMN	DEI	KRSTV	34,25
	478	N	FP	V	GIKQRWY	ACDHL	30,98
Least preferred	218	I	KNS	AHL	CDEGMV	RTW	27,86
	187	V	A	CFGQT	DEHIKLMRSWY		26,84
	215	Q	LM	CDF	AIS	GHNWY	24,63
	479	W	QST	CE	DFGHIKLMNPRVY		21,51
	222	I	D	ACH	EGMPVY	KTW	20,91
	191	A	N	DGIPT	CFHKQSVWY	ELMR	20,35
	455	W	ELMNT	PV	ACGKQSY	DFIR	19,92

The most prominent mutations were found at position 376 when the wildtype amino acid Asn was exchanged for amino acids with large hydrophobic side chains such as Phe, Ile, Leu, Met, Thr, or Val; these variants facilitated clear growth and almost full xylose consumption and almost no transport activity for glucose in the GTAC screen (Figure 8). We have found that advantageous aminoacids at position 376 have a van der Waals volume of 85 to 138 Å³ and a side chain hydrophobicity of 10 to 100 Δt_R, or in an embodiment a van der Waals volume of 100 to 138 Å³ and a side chain hydrophobicity of 10 to 100 Δt_R. A further important embodiment is specific aminoacids at position 339, where we have found an aminoacid that has a side chain hydrophobicity of -30 to 10 Δt_R results in mutants with reduced glucose transport activity. This is illustrated by figure 11.

The values for van der Waals volume ($\text{\AA}^3$) for aminoacids are herein used as described in: http://www.proteinsandproteomics.org/content/free/tables_1/table08.pdf. The corresponding literature is N. J. Darby, Thomas E. Creighton, Protein Structure (1993) Oxford University Press. The values for side chain hydrophobicity (Δt_R) of aminoacids are herein used as described in <http://onlinelibrary.wiley.com/doi/10.1002/psc.310010507/pdf>. The reference corresponding to this is Monera, O.D. et al, Journal of Peptide science Vol. 1 319-329 (1995).

Example 5: Xylose Transport Activity Screening

Aim. Using the Xylose Transport Activity (XTA) screen, *i.e.* using hexose transporter-mutant strain RN1053 as host to introduce *GAL2* variants and screening the resulting transformants on medium for growth on low xylose (1 g l^{-1}) concentrations, mutations can be identified that increase the xylose transport activity in the *Gal2p* variant. Activity of transport can be defined by more than one parameter, for instance, the affinity of the transporter (expressed by the Michaelis constant, *i.e.* K_m), or the rate of the transporter (expressed as $V_{\max}$). It is also possible that a mutation increases the expression of the transporter, and thus improves xylose transport activity in the host cell.

Transformation and colony picking. RN1053 was transformed with a total of 468 constructs, each one bearing a *GAL2* mutant. One construct with wildtype *GAL2* sequence (SEQ. ID58) was included as control. For each transformation, 1-3 colonies were re-plated to agar medium MTPs amounting to 1229 transformants. The 1229 transformants were screened in two parts. For each part of the screening, wildtype *GAL2* was included as control.

Pre-culture and screen. For the pre-culture, transformants resulting from automated transformation and colony picking were transferred by automated process from selection agar medium MTPs to 96HDWP containing in each well 250 μl Verduyn-urea supplemented with 2% maltose. After 3 days of pre-culture in an orbital shaker at 30 °C and 750 rpm, 5 μl of culture was transferred to three different 96HDWP containing 250 μl Verduyn-urea supplemented with 1 g l^{-1} xylose (RN01-medium), each 96 HDWP representing a sampling point. On each plate (24 DWP or 96 HDWP) at least one RN1001 growth control was inoculated to have an indication of plate-to-plate effects. After 24 hours, 72 hours and 96 hours a series of 96HDWPs was harvested for

automated OD-measurement at 600 nm. For each construct and each time point the OD600 values for the different replicates was averaged. In order to compare the different parts of the screen, a relative value was calculated based on the difference to the wildtype OD600 value measured in the particular part of the XTA screen, according to the following formula:

$$\text{RelOD600} = \frac{(\text{OD600}_{\text{Gal2-mutant}} - \text{OD600}_{\text{Gal2-wildtype}})}{\text{OD600}_{\text{Gal2-wildtype}}} \times 100\%$$

Results. In both parts of the XTA screen RN1001 is part of the TOP15 positions based on the growth profile compared to wildtype Gal2p; RN1001 has the full complement of hexose transporters and the presence of RN1001 in the TOP15 in both parts of the XTA Screen (Figure 9), indicates that one or more endogenous hexose transporters in yeast facilitate high affinity xylose uptake in contrast with the wildtype Gal2 RN1053-transformants which displayed poorer growth profiles. The mutant variants with the single amino acid changes present in the TOP15 displayed an increased affinity to xylose compared to wildtype Gal2p and developed towards the RN1001 growth profile or even showed similar or improved growth characteristics on low xylose concentrations. Interestingly, when aligning amino acid sequences of the *S. cerevisiae* hexose transporter family, some of the Gal2p mutations found in the TOP15, e.g. Asn346Asp and Met339Ser, are to be found in the wildtype sequences of Hxt2 (Ser324) and Hxt11 (Asp336 and Ser329), respectively. The comparison of all tested mutant *GAL2* variants versus wildtype *GAL2* on growth (average RelOD600) at 96 hours is listed in Table 8. All mutations with RelOD600>0 are proposed to have a positive effect on the xylose affinity.

Table 9. Average RelOD600 values for the screened Gal2p mutants in the XTA Screen.

Wildtype *GAL2* construct pDB1250 was set to 0. RelOD600 values are the averages of 1-3 replicates and are relative values compared to wildtype.

Pos	wtAA	MutAA	RelOD600
85	F	A	-2,50
		C	-1,81
		D	-32,03
		E	-37,55
		G	-51,06
		H	-35,08

Pos	wtAA	MutAA	RelOD600
		K	-27,75
		N	-6,14
		P	-23,27
		Q	-0,27
		R	-18,30
		S	7,30
		T	-6,07
		V	2,40
89	T	A	-1,12
		C	19,39
		E	-13,89
		I	-20,86
		K	-13,89
		L	1,12
		M	-9,48
		N	-8,70
		Q	-4,57
		R	-10,12
		V	3,11
		W	-43,47
		Y	0,01
187	V	A	-10,62
		C	14,34
		D	-5,07
		E	-2,94
		F	-6,14
		G	14,77
		H	-2,79
		I	15,55
		L	10,08
		M	36,83
		Q	26,77
		R	-11,36
		S	-28,01
		T	-10,54
		W	-17,20
		Y	11,95
191	A	C	5,10
		D	7,05
		E	0,58
		F	-1,81

Pos	wtAA	MutAA	RelOD600
		G	-4,01
		H	-12,93
		I	10,13
		K	7,11
		M	-4,82
		N	4,60
		P	1,90
		Q	-8,47
		R	3,34
		S	-19,15
		T	-2,00
		V	-9,91
		W	4,10
		Y	-15,24
214	Y	A	-3,04
		C	-12,04
		G	-6,92
		I	-6,92
		K	-4,15
		L	6,02
		M	-26,59
		N	-9,60
		P	5,35
		Q	-26,44
		R	-29,77
		S	-37,90
		W	-21,91
215	Q	A	1,54
		C	5,92
		D	25,93
		E	20,81
		F	3,28
		I	-5,50
		L	8,24
		M	11,19
		R	-48,74
		S	16,66
		V	-21,07
		W	-39,22
		Y	-24,30
218	I	A	-15,52

Pos	wtAA	MutAA	RelOD600
		C	-13,03
		D	-19,50
		E	-15,31
		G	-9,12
		H	-9,12
		K	-10,19
		L	-15,45
		M	-24,27
		N	-23,20
		R	-13,39
		S	-13,60
		T	-15,95
		V	-4,00
		W	40,30
219	T	A	-6,71
		C	-20,53
		D	-21,69
		E	3,47
		F	-17,01
		G	-35,80
		I	-47,23
		L	3,78
		M	-20,00
		N	-13,93
		P	-18,27
		Q	-14,25
		R	7,17
		S	-18,08
		V	11,19
		W	-15,13
222	I	A	-7,91
		C	-6,92
		D	-12,54
		E	-16,80
		G	28,78
		H	-13,50
		K	-48,17
		M	20,18
		P	11,01
		R	-0,62
		S	3,03

Pos	wtAA	MutAA	RelOD600
		T	10,25
		V	-12,55
		W	11,38
		Y	23,82
226	Y	A	23,92
		C	-6,78
		D	-12,89
		E	5,10
		F	-28,04
		G	-4,64
		H	-11,26
		I	-12,46
		K	-10,55
		M	24,16
		N	-4,86
		P	20,39
		Q	0,33
		R	-9,19
		S	-5,78
		T	23,80
		W	-5,71
338	Q	A	4,10
		C	10,88
		D	7,99
		E	3,59
		F	0,45
		G	5,29
		H	11,63
		I	12,95
		L	0,58
		M	4,79
		N	14,12
		P	2,21
		R	15,03
		S	-17,02
		T	-26,76
		V	-11,33
		W	-6,92
		Y	-17,30
339	M	A	28,42
		C	20,32

Pos	wtAA	MutAA	RelOD600
		D	-0,09
		E	18,97
		F	4,17
		G	-20,00
		H	-14,46
		I	-17,55
		K	-11,75
		L	21,03
		N	4,25
		P	-4,64
		Q	-3,65
		R	-7,91
		S	54,52
		T	10,79
		V	-11,90
		W	6,56
		Y	38,66
341	Q	D	-10,62
		E	-8,13
		H	2,89
		K	-21,66
		L	-22,79
		M	-11,30
		N	-13,31
		R	-10,76
		S	7,16
		T	-3,08
		V	18,89
		Y	9,37
342	Q	A	-15,66
		C	2,04
		D	-12,54
		E	-9,23
		F	33,22
		G	42,01
		H	45,21
		I	34,38
		K	12,92
		L	28,91
		M	7,99
		N	45,31

Pos	wtAA	MutAA	RelOD600
		P	-7,15
		R	44,36
		S	22,25
		T	20,62
		V	-16,07
		W	32,62
		Y	-21,41
343	L	A	15,98
		C	7,16
		D	-12,82
		E	9,65
		F	-6,07
		G	-6,63
		H	-9,83
		I	19,04
		K	2,40
		M	-0,73
		N	0,97
		P	4,53
		Q	19,53
		R	-10,62
		S	10,72
		T	-6,46
		W	0,69
		Y	29,99
346	N	A	26,36
		C	15,73
		D	69,10
		F	10,15
		G	24,01
		H	-4,93
		I	0,76
		K	-1,66
		P	-0,80
		Q	-5,28
		R	-7,84
		V	-44,97
		Y	-3,51
347	N	D	5,29
		E	-24,11
		G	7,59

Pos	wtAA	MutAA	RelOD600
		H	-4,57
		I	-3,08
		K	9,65
		L	14,52
		Q	6,55
		R	-1,93
		S	18,92
		T	2,61
		V	6,09
350	Y	D	-2,08
		E	39,52
		I	-28,19
		L	33,56
		M	7,09
		Q	6,95
		R	-40,48
		S	7,93
		T	-6,33
		V	-15,81
		Y	-9,27
373	G	A	-51,50
		E	-50,72
		F	-47,17
		H	-37,99
		I	2,84
		K	15,40
		L	18,36
		M	14,21
		N	21,06
		Q	14,02
		R	32,18
		S	18,92
		T	-11,04
		V	-6,71
		W	-5,20
		Y	1,21
376	N	A	10,19
		C	24,67
		D	8,93
		E	10,19
		F	-13,02

Pos	wtAA	MutAA	RelOD600
		H	-13,78
		I	14,59
		K	47,76
		L	-9,63
		M	6,29
		P	29,04
		Q	35,51
		R	-23,42
		S	44,42
		V	-42,40
		W	-15,74
		Y	-21,28
380	T	A	3,56
		C	8,37
		D	24,95
		E	14,21
		F	-17,33
		G	29,22
		I	20,81
		K	14,34
		L	23,45
		M	29,16
		N	-50,25
		P	-7,46
		Q	-15,32
		R	-0,55
		S	-6,02
		V	-6,21
		W	14,78
		Y	29,48
383	S	A	16,05
		C	3,32
		D	-2,94
		E	1,97
		F	-15,10
		G	-4,50
		I	2,40
		K	-47,86
		L	3,59
		M	-2,56
		N	-44,72

Pos	wtAA	MutAA	RelOD600
		Q	14,40
		R	21,37
		T	-28,26
		V	24,58
		W	26,71
		Y	27,84
444	F	C	-6,49
		D	-2,08
		E	-16,16
		G	1,54
		H	-16,38
		I	-20,29
		K	-24,06
		L	-10,19
		M	-7,20
		N	-9,76
		P	-18,79
		Q	-27,18
		S	-14,45602
		W	-16,94
		Y	-18,08
446	F	D	19,61
		E	-36,17
		H	30,98
		I	0,83
		L	-40,55
		M	-2,58
		N	-10,90
		Q	-35,54
		R	-37,62
		S	-2,94
		T	-39,41
		V	-47,79
448	T	A	10,57
		C	0,08
		D	15,78
		E	19,55
		F	-6,65
		G	16,91
		H	25,64
		I	-0,18

Pos	wtAA	MutAA	RelOD600
		K	26,65
		L	22,31
		M	0,77
		P	12,58
		Q	14,08
		R	29,22
		S	11,19
		V	
		W	8,09
		Y	-2,01
449	T	A	4,67
		C	6,59
		D	-2,58
		E	-2,08
		F	16,48
		G	11,00
		H	16,26
		I	9,51
		K	-3,51
		L	11,29
		M	-15,59
		N	-7,13
		P	8,73
		Q	13,21
		R	-10,33
		S	15,48
		V	14,41
		W	-5,78
		Y	-5,57
451	T	C	-33,23
		D	-23,20
		F	-4,75
		G	-16,94
		H	-41,76
		I	-40,26
		K	-8,91
		L	-23,73
		N	-6,07
		P	-6,14
		Q	-8,22
		R	-6,21

Pos	wtAA	MutAA	RelOD600
		S	2,90
		T	-11,42
		V	-5,89
		W	4,10
		Y	3,91
455	W	C	-14,75
		D	8,18
		E	4,22
		F	-2,31
		G	-10,67
		I	5,45
		K	8,12
		L	-5,96
		M	6,39
		N	-9,22
		P	0,83
		Q	0,95
		R	-48,55
		S	-4,82
		T	-9,44
		V	-2,06
		Y	-19,81
478	N	A	-25,74
		C	-14,75
		D	-9,91
		F	6,86
		G	0,01
		H	-13,12
		I	8,81
		K	-24,30
		P	-20,47
		Q	14,40
		R	-27,44
		W	-13,40
		Y	6,77
479	W	C	-12,96
		D	15,20
		E	18,40
		F	24,87
		H	1,83
		I	48,83

Pos	wtAA	MutAA	RelOD600
		K	8,58
		L	-3,79
		M	7,87
		N	8,05
		P	-42,40
		Q	-33,73
		R	13,42
		S	57,22
		T	15,48
		V	-12,54
		Y	12,07

Table 10. Gal2p positions and mutations identified in the XTA screening are ordered on preference. SCORE TOP HIT is the RelXyl score for the amino acid substitution within a position yielding the clearest improvement compared to wildtype Gal2p This allows a sorting of the most influential mutations and relevant positions to target in Gal2p to increase xylose transport activity.

	XTA	Mutations					
	Position	wt AA	Most preferred	Preferred	Least preferred	Inactive	SCORE TOP HIT
Most preferred	346	N	D	AG	CF	HIKPQRVY	69,10
	479	W	S	I	DEFHKMNR TY	CLPQV	57,22
	339	M	SY	ACEL	FNTW	DGHIKPQRV	54,52
	342	Q	GHNR	FILW	CKMST	ADEPVY	47,76
	376	N	KS	CPQ	ADEIM	FHLRVWY	47,76
Preferred	218	I	W			ACDEGHKLMNRS TV	40,30
	350	F	E	L	MQS	DIRTVY	39,52
	187	V	MQ	CGI	LY	ADEFHRSTW	36,83
	373	G	R	KLMNQ S	IY	AEFHTVW	32,18
	446	F	H	D		EILMNQRSTV	30,98

	XTA	Mutations					
Least preferred	343	L	Y	AIQ	CEKSP	DFGHMNRTW	29,99
	380	T	GMY	DIL	ACEKW	FNPQRSV	29,48
	448	T	HKR	EL	ADGPQSW	MCFIVY	29,22
	222	I	G	MY	PTW	ACDEHKRSV	28,78
	383	S	VWY	AQR	CEIL	DFGKMNT	27,84
	215	Q	DE	S	ACFL	IMRVWY	25,93
	226	Y	AMPT		E	CDFGHIKNQRSW	24,16
	89	T	C		LVY	AEIKMNQRW	19,39
	347	N	S	KL	DGQTV	EHIR	18,92
	449	T	FH	QSV	ACGILP	DEKMNRWY	16,48
	338	Q	NR	CDHI	AEFGLMP	STVWY	15,03
	478	N	Q	FIY		ACDGHKPRW	14,40
	191	A	I	DK	CENPRW	FGHMQSTVY	10,13
	341	Q	V	SY	H	DEKLMNRT	9,37
	455	W	DK	EIM		CFGLNPQRSTVY	8,18
	85	F	S	V		ACDEGHKNPQRT	7,30
	219	T	V	R	EL	ACDFGIMNPQSW	7,17
	214	Y	LP			ACGIKMNQRSW	6,02
	451	A	WY	S		CDFGHIKLNQRT V	4,10
	444	F			G	CDEHIKLMNPQS WY	1,54

REFERENCES

Ausubel *et al.* 1995 Current Protocols in Molecular Biology, John Wiley & Sons, Inc.

5 Becker J, Boles E. 2003 A modified *Saccharomyces cerevisiae* strain that consumes L-Arabinose and produces ethanol. *Appl Environ Microbiol.* 69 p. 4144-4150.

 Hamacher T, Becker J, Gárdonyi M, Hahn-Hägerdal B, Boles E. 2002 Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization. *Microbiology* 148 p. 2783-2788.

10 Kasahara T, Kasahara M. 2000 Three aromatic amino acid residues critical for galactose transport in yeast Gal2 transporter. *J Biol Chem.* 275 p. 4422-4428.

 Kasahara T, Maeda M, Boles E, Kasahara M. 2009 Identification of a key residue determining substrate affinity in the human glucose transporter GLUT1. *Biochim Biophys Acta.* 1788 p.1051-1055.

15 Kasahara T, Kasahara M. 2010 Identification of a key residue determining substrate affinity in the yeast glucose transporter Hxt7: a two-dimensional comprehensive study. *J Biol Chem.* 285 p. 26263-26268.

 Kuyper M, Hartog MM, Toirkens MJ, Almering MJ, Winkler AA, van Dijken JP, Pronk JT 2005 Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. *FEMS Yeast Research* 5 p. 399-409.

 Luttik MA, Kötter P, Salomons FA, van der Klei IJ, van Dijken JP, Pronk JT 2000 The *Saccharomyces cerevisiae* *ICL2* gene encodes a mitochondrial 2-methylisocitrate lyase involved in propionyl-coenzyme A metabolism. *Journal of Bacteriology* 182 p.7007-7013.

25 Nelissen B, De Wachter R, Goffeau A. 1997 Classification of all putative permeases and other membrane plurispanners of the major facilitator superfamily encoded by the complete genome of *Saccharomyces cerevisiae*. *FEMS Microbiol Rev.* 21 p. 113-134.

30 Sambrook *et al.* 1989 Molecular Cloning, a Laboratory Manual

 Schiestl RH and Gietz RD 1989 High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Current Genetics* 16 p.339-346.

Wieczorke R, Krampe S, Weierstall T, Freidel K, Hollenberg CP, Boles E 1999
Concurrent knock-out of at least 20 transporter genes is required to block uptake of
hexoses in *Saccharomyces cerevisiae*. *FEBS Letters* 464 p.123-128.

5 Young E, Poucher A, Comer A, Bailey A, Alper H. 2011 Functional survey for
heterologous sugar transport proteins, using *Saccharomyces cerevisiae* as a host. *Appl*
Environ Microbiol. 77 p. 3311-3319.

CLAIMS

1. Polypeptide having one or more substitution at a position corresponding to position 339 or 376 of SEQ ID NO: 59, wherein the polypeptide is a member of the Major Facilitator Superfamily (MFS).
5
2. Polypeptide according to claim 1, wherein the substitution is at position corresponding to 376 and wherein the aminoacid at that position is replaced by an aminoacid that has a van der Waals volume of 80 to 138 Å³ and a side chain hydrophobicity of 10 to 100 Δt_R.
10
3. Polypeptide according to claim 1, wherein the substitution is at position corresponding to 339 and wherein the aminoacid at that position is replaced by an aminoacid that has a side chain hydrophobicity of -30 to 10 Δt_R.
15
4. Polypeptide according to any of claims 1 to 3, having one or more aminoacid corresponding to 339N/V, or 376I/M/V.
5. Polypeptide according to claim 1, wherein the polypeptide has reduced glucose transport activity compared to the polypeptide of SEQ ID NO: 59.
20
6. Polypeptide according to claim 2, wherein the polypeptide has improved xylose transport activity compared to the polypeptide of SED ID NO: 59.
7. Polypeptide according to any of claims 1 to 3, wherein the polypeptide has increased reduced glucose transport activity and improved xylose transport activity compared to the polypeptide having SEQ ID NO: 59.
25
8. Polypeptide according to any of claims 1 to 7, comprising sequences that contain one or more of the following amino acid motifs:
30
 - a) G-R-x(3)-G-x(3)-G-x(11)-E-x(5)-[LIVM]-R-G-x(12)-[GA] ;
 - b) R-x(14)-G-x(2)-Y-x(2)-[YF]-[YF]-[GSAL] and/or
 - c) V-x(15)-[GNR]-[RH]-R-x(2)-[LM]-x(2)-[GA]
 wherein the fragments in (a) to (c) in the polypeptide correspond to the positions in

SEQ ID NO: 59.

9. Polypeptide according to any of claims 1 to 8 comprising a motif G-R-x(3)-G-x(3)-G-x(11)-E-x(5)-[LIVM]-R-G-x(12)-[GA].

5

10. Polynucleotide having at least 50% identity to SEQ ID NO: 56, encoding the polypeptide according to any of claims 1 to 9.

11. Nucleic acid construct comprising the polynucleotide of claim 10.

10

12. Host cell transformed with the nucleic acid construct of claim 11.

13. Transformed host cell according to claim 12, which is an eukaryote.

15

14. Transformed host cell according to claim 13, which is yeast.

15. Transformed host cell according to claim 14, which belongs to the genus *Saccharomyces*, in particular the species *Saccharomyces cerevisiae*.

20

16. Transformed host cell comprising a heterologous nucleotide that encodes a polypeptide according to any one of claims 1 to 9 or encodes a polypeptide having substitution M339S or N376C of sequence ID NO: 59.

25

17. Transformed host according to claim 16, where in the polynucleotide encodes a polypeptide that is a mutant of a polypeptide that is native in the untransformed host cell.

30

18. Transformed host cell according to claim 16 or 17, wherein the polypeptide that is native in the untransformed host cell is a member of the the Major Facilitator Superfamily (MFS).

19. Transformed host cell according to any of claims 17 or 18, wherein the polypeptide that is native in the untransformed host cell is a hexose transporter polypeptide.

20. Transformed host cell according to any of claims 17 to 19, wherein the polypeptide that is native in the untransformed host cell is a hexose transporter polypeptide.
21. Transformed host cell according to claim 20, wherein the polypeptide that is native
5 in the untransformed host cell is a transporter polypeptide chosen from the list consisting of Gal2, Hxt1, Hxt2, Hxt3, Hxt4, Hxt5, Hxt6, Hxt7, Hxt8, Hxt9, Hxt10, Hxt11, Hxt12, Hxt13, Hxt14, Hxt15, Hxt16an Hxt17.
22. Transformed host cell that when subjected to the GATC protocol, consumes xylose
10 from a medium comprising xylose and glucose, while glucose is still present in the medium.
23. Transformed host cell according to claim 22 that that when subjected to the GATC protocol consumes xylose faster than it consumes glucose.
- 15 24. Use of a polypeptide having substitution corresponding to N376T of SEQ ID NO:59, as a xylose transporter, wherein the polypeptide is a member of the Major Facilitator Superfamily (MFS).
- 20 25. Use according to claim 22, wherein the polypeptide is expressed in an eukaryotic cell and the eukaryotic cell is used to ferment xylose in the presence of glucose.
26. Process for the degradation of ligno-cellulosic or hemi-cellulosic material, wherein ligno-cellulosic or hemi-cellulosic material is contacted with an enzyme
25 composition, wherein one or more sugar is produced, and wherein the produced sugar is fermented to give a fermentation product, wherein the fermentation is conducted with a transformed host cell of any of claims 12 to 23.
27. Process according to claim 26, wherein the produced sugar comprises xylose and
30 glucose and wherein the cell co-ferments xylose and glucose.
28. Process according to claim 27, wherein the fermentation product is one or more of ethanol, butanol, lactic acid, a plastic, an organic acid, a solvent, an animal feed supplement, a pharmaceutical, a vitamin, an amino acid, an enzyme or a chemical

feedstock.

29. Process for the degradation of ligno-cellulosic or hemi-cellulosic material, wherein ligno-cellulosic or hemi-cellulosic material is contacted with an enzyme composition, wherein one or more sugar is produced, and wherein the produced sugar is fermented to give a fermentation product, wherein the fermentation is conducted with a transformed host cell of any of claims 12 to 23.
30. Process according to claim 29, wherein the fermentation product is one or more of ethanol, butanol, lactic acid, di-terpene, glycosylated di-terpene, a plastic, an organic acid, a solvent, an animal feed supplement, a pharmaceutical, a vitamin, an amino acid, an enzyme or a chemical feedstock.

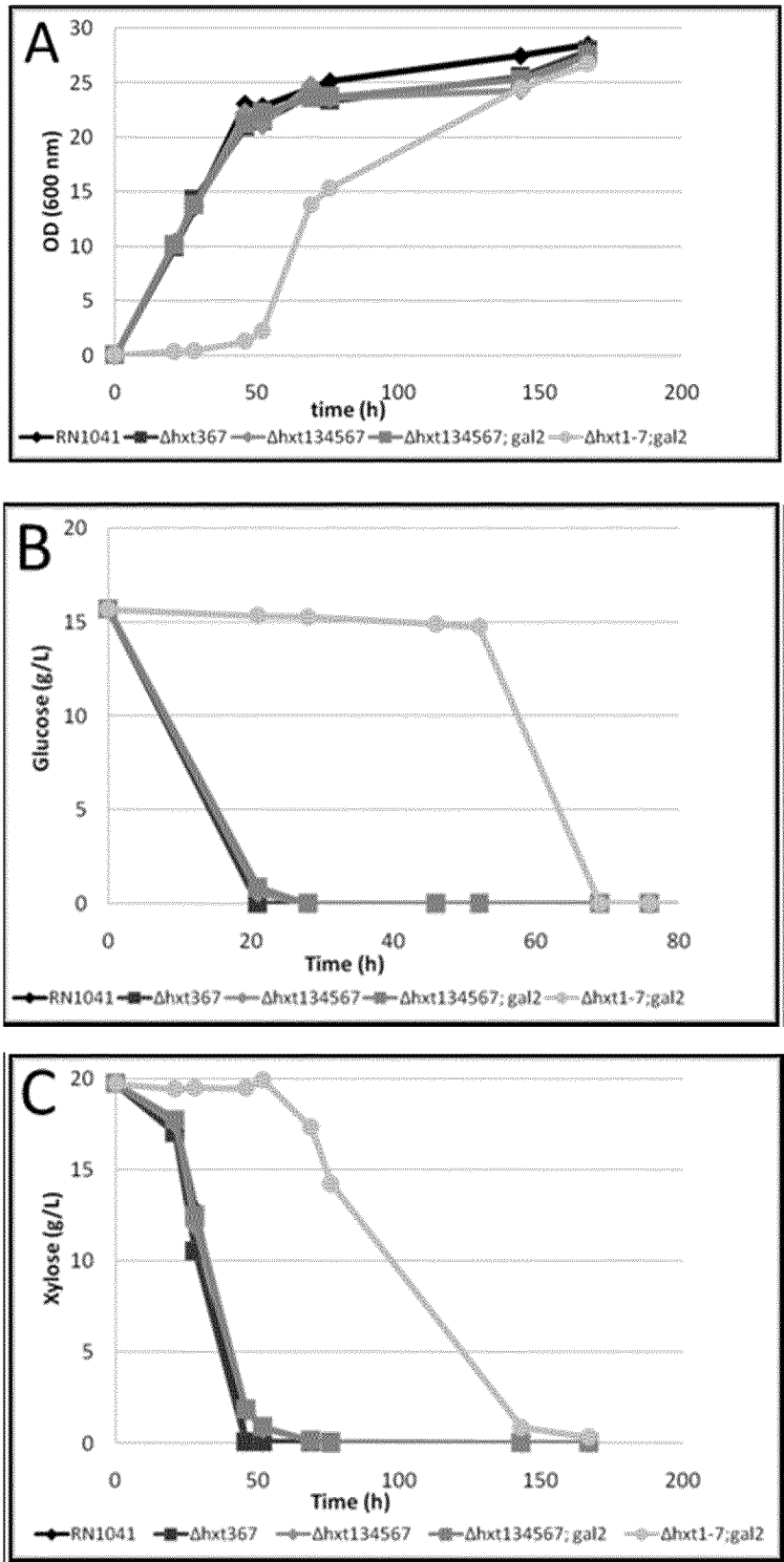


FIG. 1

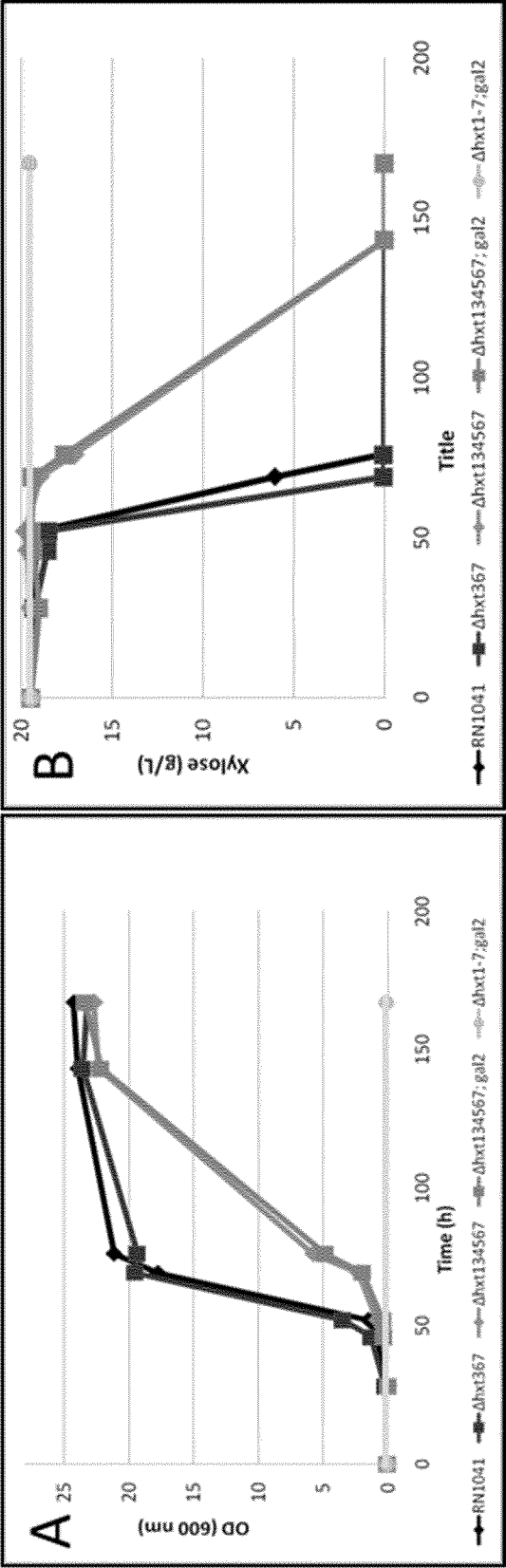


FIG. 2

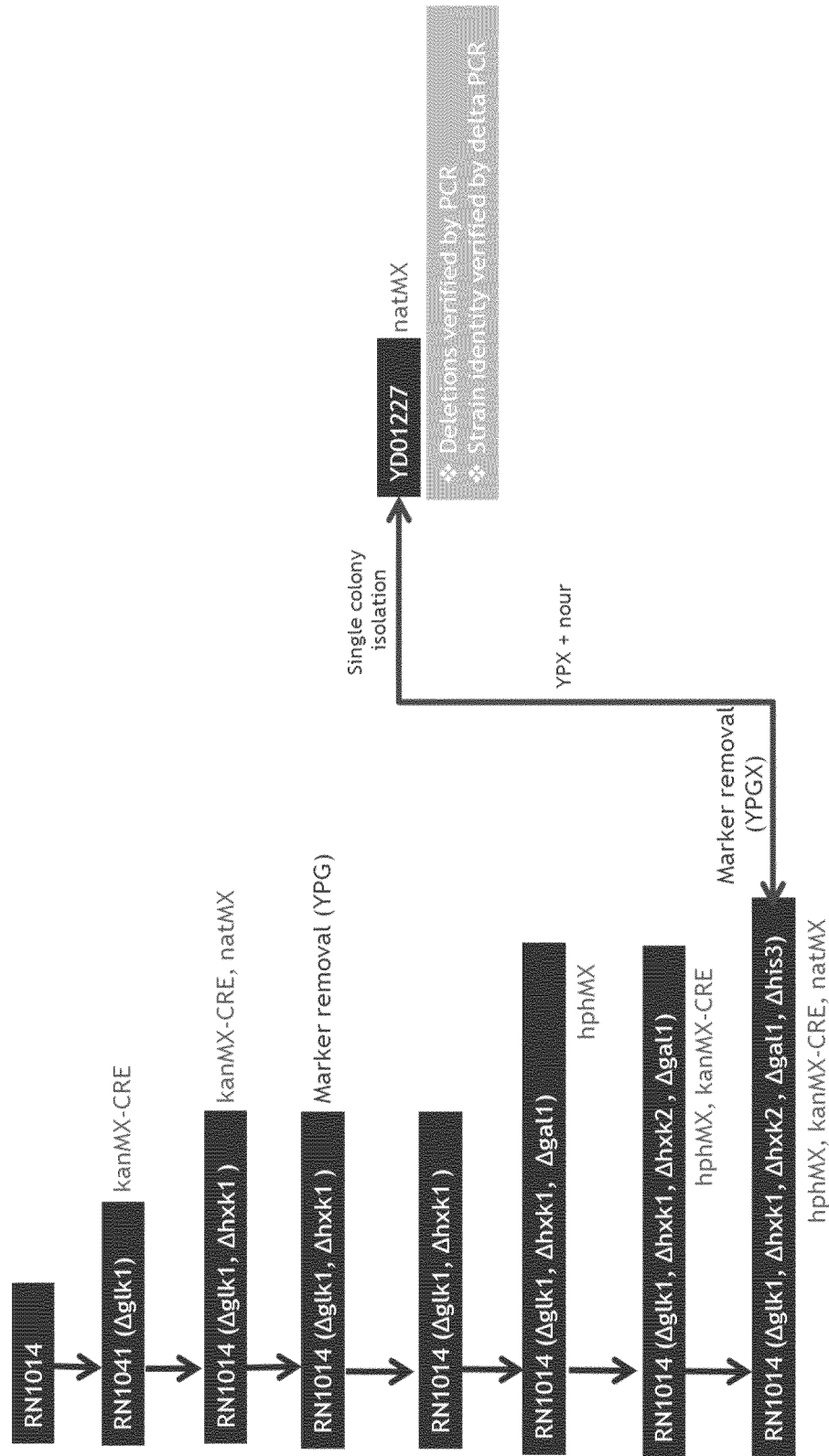


FIG. 3

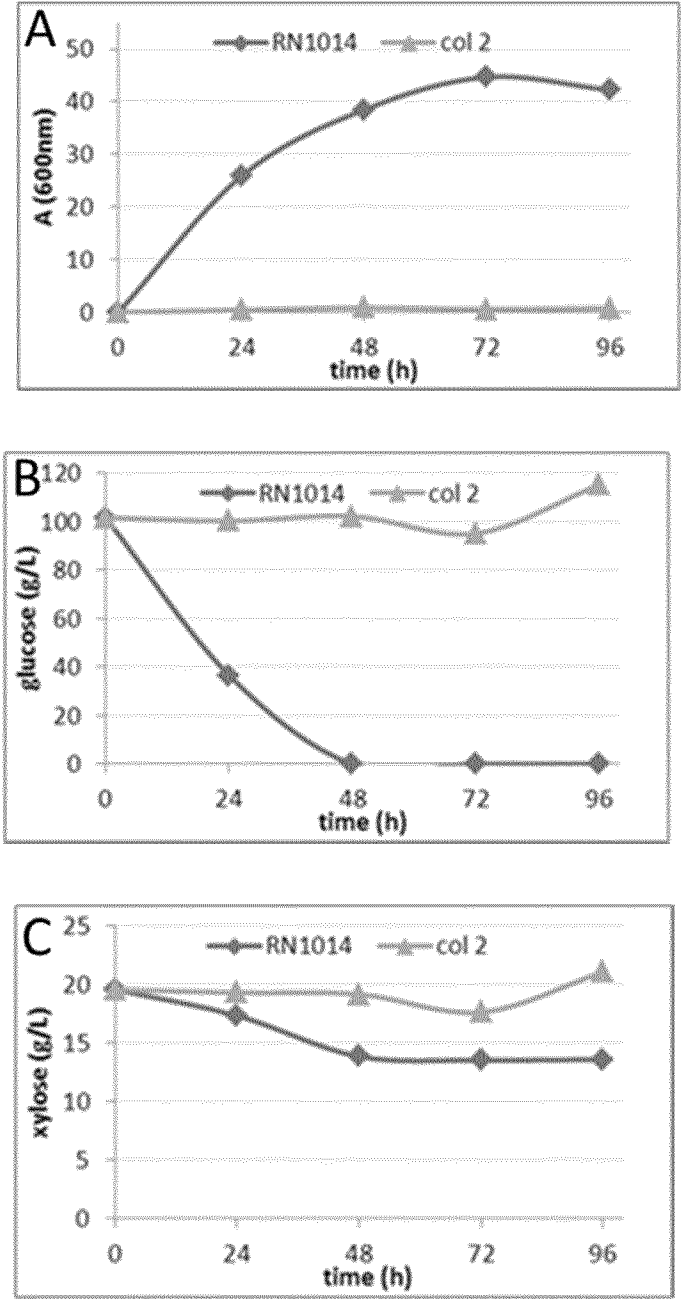


FIG. 4

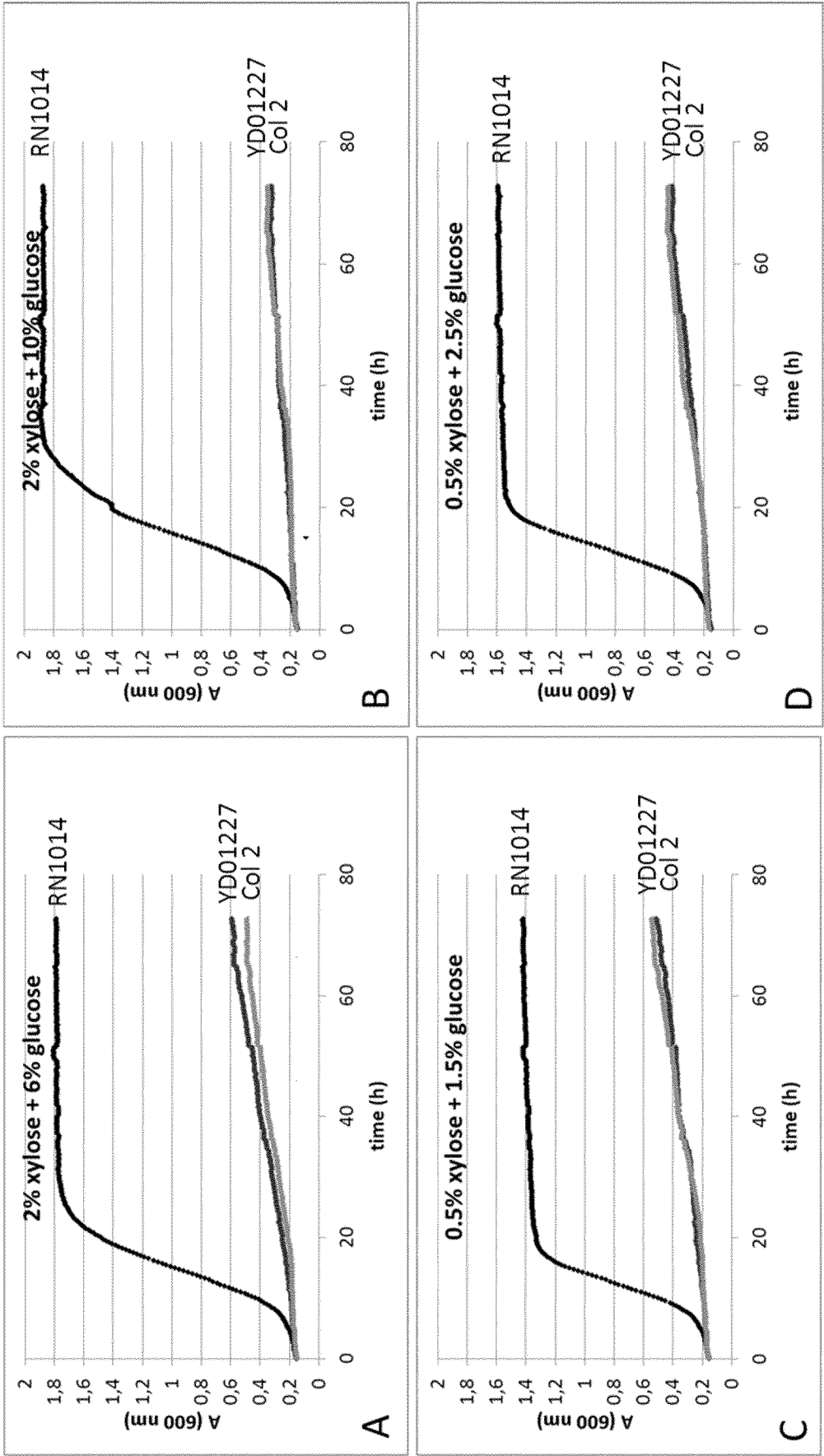


FIG. 5

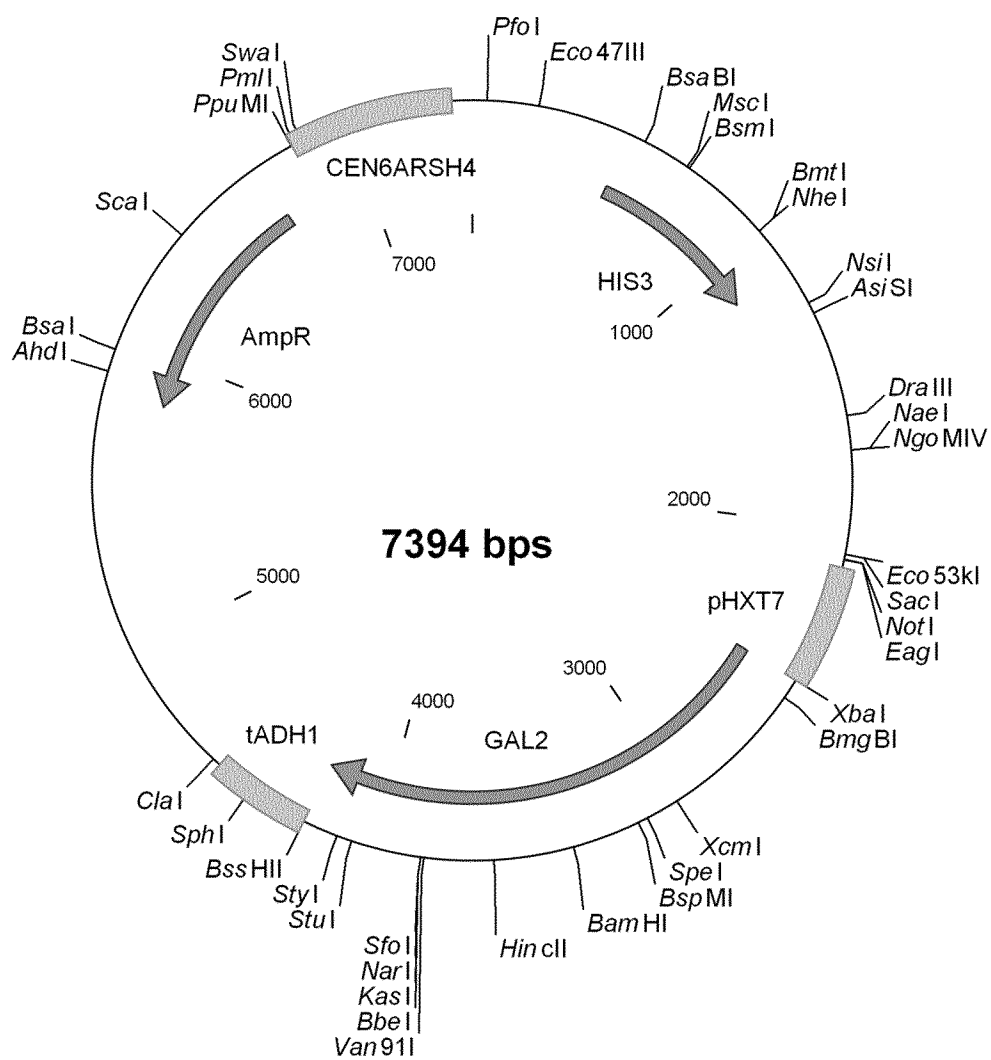


FIG. 6

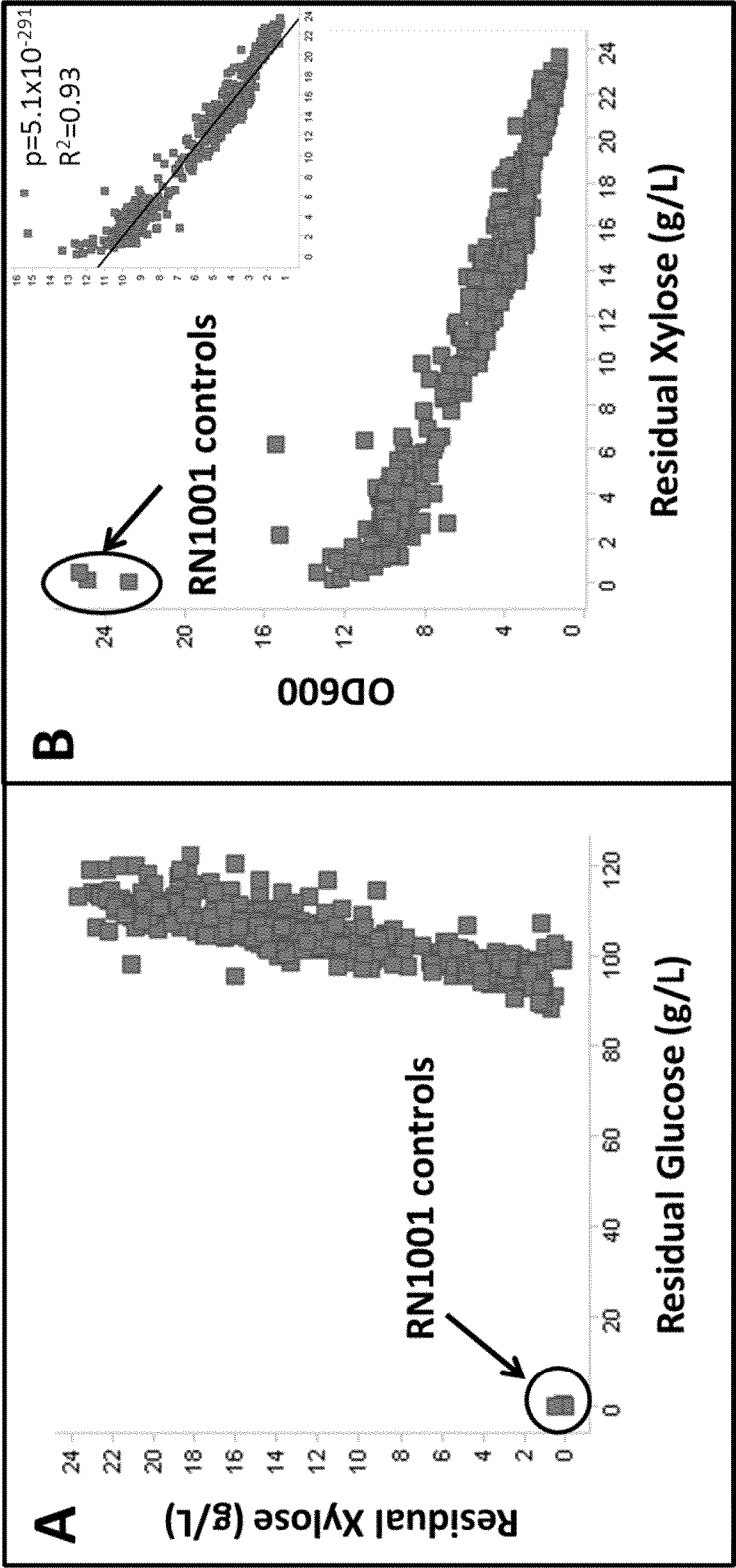


FIG. 7

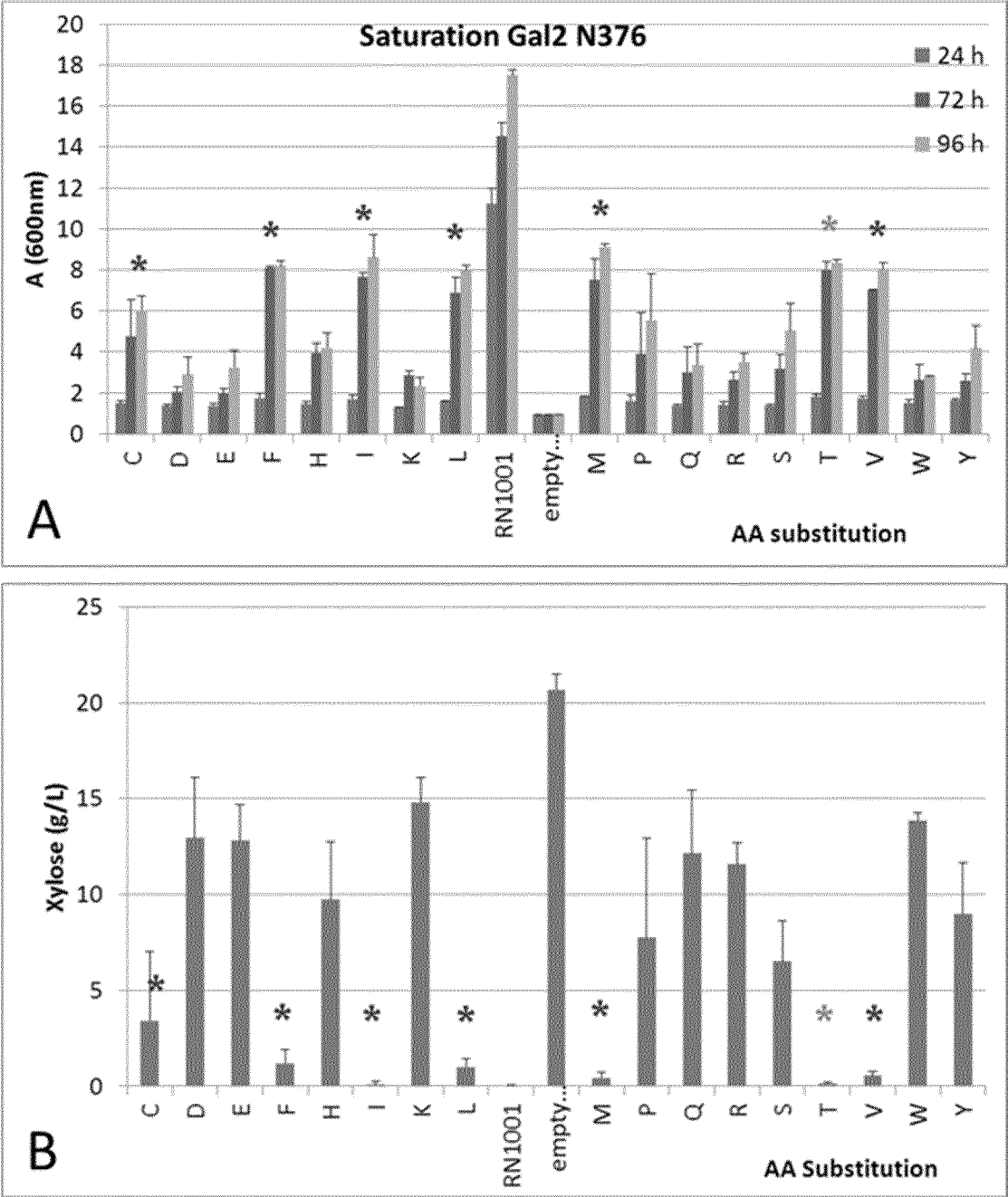


FIG. 8

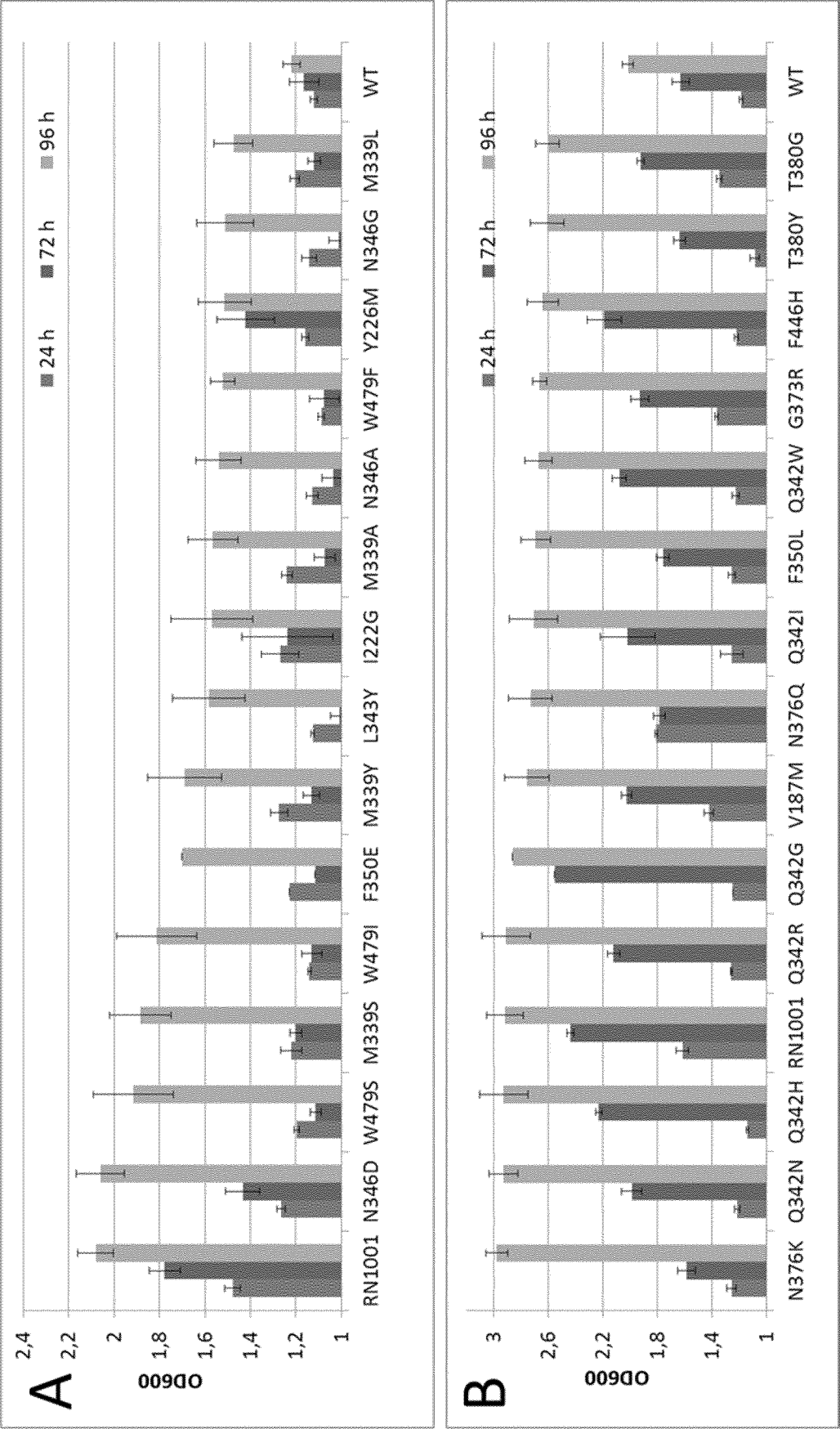
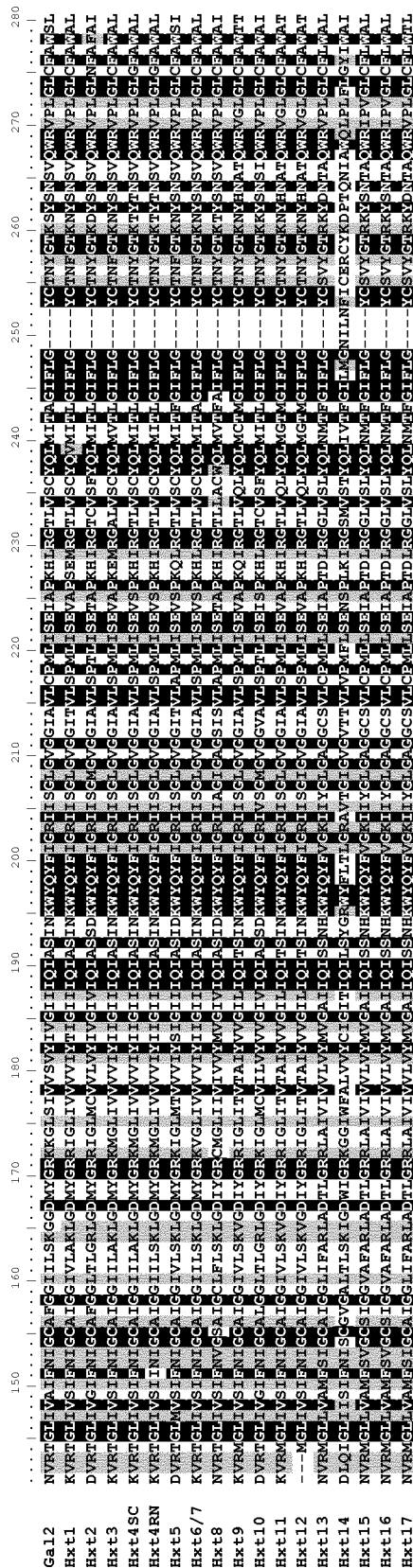


FIG. 9



Gal2	LWTVENLGRKCLLLCAATMACAVIVASVGVTRLPAG--KQSPSSKAGACMIVTCFYIFCEATWAPVAVNIIVAESPELVRKSKMALASNAWNGELIAFPPTPTTSAINWYGVVFGCLVAMFFVVFVFFVE	430
Hxt1	LYTVDFGRNCLLWGAAGVCCLVVIVASVGVTRLPAG--QNNSSKAGACMIVTCFYIFCEATWAPIAVVISECEPLVRKSKMSIASAANWNGELISFPPTPTGAINWYGVVFGCLWFAFYVVFVFFVE	440
Hxt2	LYTVDFGRKCLLGSASMAICVIVSTVGVTRLPAG--KQSPSSKAAGNMIVTCFLIFCEATSWAPIAVVIVAESPELVRKKNAMALAVGANWNGELIGFTPTPTTSALGWSGVVFGCLVFSFVVFVFFVQ	450
Hxt3	LYTVDFGRNCLLWGAAGVCCLVVIVASVGVTRLPAG--EGNSSKAGACMIVTCFYIFCEATWAPIAVVISEPELVRKSKMSIATAANWNGELIGFTPTPTGAINWYGVVFGCLWFAFYVVFVFFVE	460
Hxt4	IFLVERYGRKCLLWGAASMAICVIVASVGVTRLPAG--KKNSSKAGACMIVTCFYIFCEATWAPIPVFWNSETPELVRKSKMALAQACWNGELIGFTPTPTSCALDRYGVVFGCLVFSFYVVFVFFVE	470
Hxt4S	IFLVERYGRKCLLWGAASMAICVIVASVGVTRLPAG--KKNSSKAGACMIVTCFYIFCEATWAPIPVFWNSETPELVRKSKMALAQACWNGELIGFTPTPTSCALDRYGVVFGCLVFSFYVVFVFFVE	480
Hxt4RN	LYTVDFGRNCLLWCVGMICLVVIVASVGVTRLPAG--QDQPSKAGACMIVTCFYIFCEATWAPVAVVISEPELVRVRGRAMSIASACWNGELISFPPTPTTSAINWYGVVFGCLWFAFYVVFVFFVE	490
Hxt5	LYTVDFGRNCLLWGAAGVCCLVVIVASVGVTRLPAG--QDQPSKAGACMIVTCFYIFCEATWAPIPVVIVSETPELVRKSKMSIATAANWNGELIGFTPTPTGAINWYGVVFGCLVFMFFVVFVFFVE	500
Hxt6/7	LYTVDKLGRKCLLCAATMACAVIVASVGVTRLPAG--KSEPSKAGACMIVTCFYIFCEATWAPVAVVISEPELVRKSKMSIATAANWNGELIGFTPTPTTSAINWYGVVFGCLVFSFYVVFVFFVE	510
Hxt8	VYTIERFGRKCLLWGAASMLCFAFVASVGVTRLPAGSSHQDITISQAGACMIVTCFYIFCEATWAGGCVVIVSETPELVRKSRGMALATAANWNGELISFPPTPTGAINWYGVVFGCLVFAFYVVFVFFVE	520
Hxt9	LYTVDFGRKCLLWGSASMAICVIVATVGVTRLPAG--KQDQSSQAGNMIVTCFFIFSEATWAPIAVVIVAEPELVRKKNAMALAVGANWNGELIGFTPTPTTSALGWSGVVFGCLVFSFYVVFVFFVE	530
Hxt10	VYTIERFGRKCLLWGAASMLCFAFVASVGVTRLPAGSSHQDITISQAGACMIVTCFYIFCEATWAGGCVVIVSETPELVRKSRGMALATAANWNGELISFPPTPTGAINWYGVVFGCLVFAFYVVFVFFVE	540
Hxt11	VYTIERFGRKCLLWGAASMLCFAFVASVGVTRLPAGSSHQDITISQAGACMIVTCFYIFCEATWAGGCVVIVSETPELVRKSRGMALATAANWNGELISFPPTPTGAINWYGVVFGCLVFAFYVVFVFFVE	550
Hxt12	VYTIERFGRKCLLWGAASMLCFAFVASVGVTRLPAGSSHQDITISQAGACMIVTCFYIFCEATWAPVAVVIVAESPEPSKVRKMSISTACWNGELIGFTPTPTGAINWYGVVFGCLVAMFFVVFVFFVE	
Hxt13	VMVVDKIGRKKLLFAAGMACAVIVASICVKCLYPHC--QDQSPSSKAGNMIVTCFYIFCEATWAPVAVVIVAESPEPSKVRKMSISTACWNGELIGFTPTPTGAINWYGVVFGCLVAMFFVVFVFFVE	
Hxt14	IYYVEKWKHTCLLYGKSTNLLFYAMTVATGCT-----FGRETFDSNIVLLIVTCCFIDREALILGPVTVIVYSELPELRPAISMAICTFFIMANELLILITPIMVSKDKELGYPAAQLALITFSMLIVE	
Hxt15	VMVVDKIGRKKLLFAASMACAVIVASICVKCLYPHC--QDQSPSSKAGNMIVTCFYIFCEATWAPVAVVIVAESPEPSKVRKMSISTACWNGELIGFTPTPTGAINWYGVVFGCLVAMFFVVFVFFVE	
Hxt16	VMVVDKIGRKKLLFAASMACAVIVASICVKCLYPHC--QDQSPSSKAGNMIVTCFYIFCEATWAPVAVVIVAESPEPSKVRKMSISTACWNGELIGFTPTPTGAINWYGVVFGCLVAMFFVVFVFFVE	
Hxt17	VMVVDKIGRKKLLFAAGMACAVIVASICVKCLYPHC--QDQSPSSKAGNMIVTCFYIFCEATWAPVAVVIVAESPEPSKVRKMSISTACWNGELIGFTPTPTGAINWYGVVFGCLVAMFFVVFVFFVE	

Gal2	LWTVENLGRKCLLLCAATMACAVIVASVGVTRLPAG--KQSPSSKAGACMIVTCFYIFCEATWAPVAVNIIVAESPELVRKSKMALASNAWNGELIAFPPTPTTSAINWYGVVFGCLVAMFFVVFVFFVE	430
Hxt1	LYTVDFGRNCLLWGAAGVCCLVVIVASVGVTRLPAG--QNNSSKAGACMIVTCFYIFCEATWAPIAVVISECEPLVRKSKMSIASAANWNGELISFPPTPTGAINWYGVVFGCLWFAFYVVFVFFVE	440
Hxt2	LYTVDFGRKCLLGSASMAICVIVSTVGVTRLPAG--KQSPSSKAAGNMIVTCFLIFCEATSWAPIAVVIVAESPELVRKKNAMALAVGANWNGELIGFTPTPTTSALGWSGVVFGCLVFSFVVFVFFVQ	450
Hxt3	LYTVDFGRNCLLWGAAGVCCLVVIVASVGVTRLPAG--EGNSSKAGACMIVTCFYIFCEATWAPIAVVISEPELVRKSKMSIATAANWNGELIGFTPTPTGAINWYGVVFGCLWFAFYVVFVFFVE	460
Hxt4	IFLVERYGRKCLLWGAASMAICVIVASVGVTRLPAG--KKNSSKAGACMIVTCFYIFCEATWAPIPVFWNSETPELVRKSKMALAQACWNGELIGFTPTPTSCALDRYGVVFGCLVFSFYVVFVFFVE	470
Hxt4S	IFLVERYGRKCLLWGAASMAICVIVASVGVTRLPAG--KKNSSKAGACMIVTCFYIFCEATWAPIPVFWNSETPELVRKSKMALAQACWNGELIGFTPTPTSCALDRYGVVFGCLVFSFYVVFVFFVE	480
Hxt4RN	LYTVDFGRNCLLWCVGMICLVVIVASVGVTRLPAG--QDQPSKAGACMIVTCFYIFCEATWAPVAVVISEPELVRVRGRAMSIASACWNGELISFPPTPTTSAINWYGVVFGCLWFAFYVVFVFFVE	490
Hxt5	LYTVDFGRNCLLWGAAGVCCLVVIVASVGVTRLPAG--QDQPSKAGACMIVTCFYIFCEATWAPIPVVIVSETPELVRKSKMSIATAANWNGELIGFTPTPTGAINWYGVVFGCLVFMFFVVFVFFVE	500
Hxt6/7	LYTVDKLGRKCLLCAATMACAVIVASVGVTRLPAG--KSEPSKAGACMIVTCFYIFCEATWAPVAVVISEPELVRKSKMSIATAANWNGELIGFTPTPTTSAINWYGVVFGCLVFSFYVVFVFFVE	510
Hxt8	VYTIERFGRKCLLWGAASMLCFAFVASVGVTRLPAGSSHQDITISQAGACMIVTCFYIFCEATWAGGCVVIVSETPELVRKSRGMALATAANWNGELISFPPTPTGAINWYGVVFGCLVFAFYVVFVFFVE	520
Hxt9	LYTVDFGRKCLLWGSASMAICVIVATVGVTRLPAG--KQDQSSQAGNMIVTCFFIFSEATWAPIAVVIVAEPELVRKKNAMALAVGANWNGELIGFTPTPTTSALGWSGVVFGCLVFSFYVVFVFFVE	530
Hxt10	VYTIERFGRKCLLWGAASMLCFAFVASVGVTRLPAGSSHQDITISQAGACMIVTCFYIFCEATWAGGCVVIVSETPELVRKSRGMALATAANWNGELISFPPTPTGAINWYGVVFGCLVFAFYVVFVFFVE	540
Hxt11	VYTIERFGRKCLLWGAASMLCFAFVASVGVTRLPAGSSHQDITISQAGACMIVTCFYIFCEATWAGGCVVIVSETPELVRKSRGMALATAANWNGELISFPPTPTGAINWYGVVFGCLVFAFYVVFVFFVE	550
Hxt12	VYTIERFGRKCLLWGAASMLCFAFVASVGVTRLPAGSSHQDITISQAGACMIVTCFYIFCEATWAPVAVVIVAESPEPSKVRKMSISTACWNGELIGFTPTPTGAINWYGVVFGCLVAMFFVVFVFFVE	
Hxt13	VAVVDKIGRKKCLLFAAGMACAVIVASICVKCLYPHC--QDQSPSKAGNMIVTCFYIFCEATWAPVAVVIVAESPEPSKVRKMSISTACWNGELIGFTPTPTGAINWYGVVFGCLVAMFFVVFVFFVE	
Hxt14	IYYVEKWKHTCLLYESTNLLFYAMTVATVGT-----FGRETFDSNIVLLIVTCCFIDREALILGPVTVIVYSELPELRPAISMAICTFFIMANELLILITPIMVSKDKELGYPAAQLALITFSMLIVE	
Hxt15	VAVVDKIGRKKCLLFAASMACAVIVASICVKCLYPHC--QDQSPSKAGNMIVTCFYIFCEATWAPVAVVIVAESPEPSKVRKMSISTACWNGELIGFTPTPTGAINWYGVVFGCLVAMFFVVFVFFVE	
Hxt16	VAVVDKIGRKKCLLFAASMACAVIVASICVKCLYPHC--QDQSPSKAGNMIVTCFYIFCEATWAPVAVVIVAESPEPSKVRKMSISTACWNGELIGFTPTPTGAINWYGVVFGCLVAMFFVVFVFFVE	
Hxt17	VAVVDKIGRKKCLLFAAGMACAVIVASICVKCLYPHC--QDQSPSKAGNMIVTCFYIFCEATWAPVAVVIVAESPEPSKVRKMSISTACWNGELIGFTPTPTGAINWYGVVFGCLVAMFFVVFVFFVE	

```

      570      580      590      600      610      620
      |.....|.....|.....|.....|.....|.....|.....|
TKGLSLEEIQELWEEGVLPWKSEGVIPSSRRGNNYDLEDLQHDDKPWYKAMLE-----
TKGLSLEEVNDMYAEGVLPWKASWVPVSKRGADYNADLMHDDQPFYKSLFSRK-----
TKGLTLEE VNEMYVEGVKPKSGSWISKEKRVSEE-----
TKGLTLEE VNEMYAEGVLPWKASWVPTSQRGANYDADALMHDDQPFYKKMF GK-----
TKGLTLEE VNTLWEEGVLPWKSPSFHQTR---EVLTTTLM-----
TKGLTLEE VNTLWEEGVLPWKSPSWPPNKRGTYNADLMHDSGTHFTRRCSEKSRSVN
TKGLTLEE VNEMYEENVLPWKSTKIIPSSRRTTDYDLDA TRNDPRPFYKRMFTKEK-----
TKGLTLEE VNTMWEEGVLPWKASWVPPSRRGANYDAEEMTHDDKPLYKRMFSTK-----
TKGLTLEE VDEMMDGVLPWKSESWVPASRRDGDYDNEKLQHDEKPFYKRMF-----
TKGLTLEE VNTMWLEGVPAWKASWVPPERRTADYDADAIDHDDRPIYKRFFSS-----
TKGLTLEE VNEMYEERIKPKSGGWIIPSSRRTPQTSSTPLVIVDSK-----
TKGLTLEE VNTMWLEGVPAWKASWVPPERRTADYDADAIDHNRPIYKRFFSS-----
TKGLTLEE VNTMWLEGVPAWKASWVPPERRTADYDADAIDHNRPIYKRFFSS-----
TIGLSLEEIQLLYEEGIKPKASWVPPSRRGISSEESKTEKKDWKKFLKFSKNSD-----
TRKKNEQEIINKIFEPE-----
TIGLSLEEIQLLYEEGIKPKASWVPPSRRGASSRETEAKKKSWEVLKFPKSFN-----
TIGLSLEETQLLYEEGIKPKASWVPPSRRGASSRETEAKKKSWEVLKFPKSFN-----
TIGLSLEEIQLLYEEGIKPKASWVPPSRRGIPSEESKTEKKDWKKFLKFSKGS D-----
SC
RN
/ 7
0
1
2
3
4
5
6
7

```

FIG. 10C

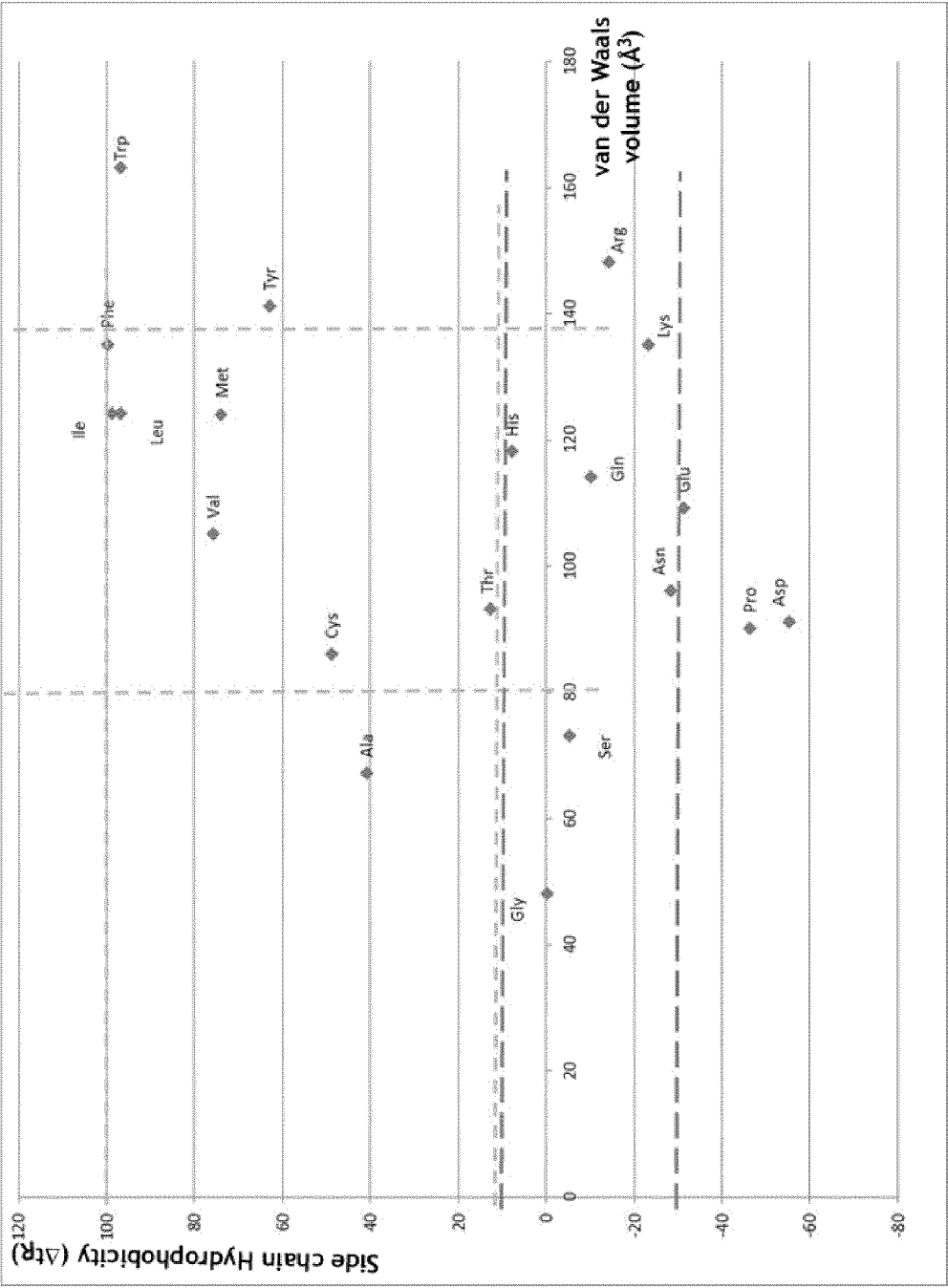


FIG. 11