



(22) **Date de dépôt/Filing Date:** 2005/07/15
 (41) **Mise à la disp. pub./Open to Public Insp.:** 2006/01/26
 (62) **Demande originale/Original Application:** 2 573 483
 (30) **Priorités/Priorities:** 2004/07/16 (US60/588,381);
 2004/07/16 (EP04077073.7); 2004/07/22 (US60/589,833)

(51) **Cl.Int./Int.Cl. C12N 1/19** (2006.01),
C12N 1/15 (2006.01), **C12N 15/53** (2006.01),
C12N 15/54 (2006.01), **C12N 15/61** (2006.01),
C12P 1/02 (2006.01), **C12P 13/00** (2006.01),
C12P 17/00 (2006.01), **C12P 19/24** (2006.01),
C12P 7/02 (2006.01), **C12P 7/06** (2006.01),
C12P 7/40 (2006.01)

(71) **Demandeur/Applicant:**
 TECHNISCHE UNIVERSITEIT DELFT, NL

(72) **Inventeurs/Inventors:**
 WINKLER, AARON ADRIAAN, NL;
 KUYPER, SIPKO MAARTEN, NL;

(54) **Titre : GENIE METABOLIQUE DE CELLULES EUCARYOTES DE FERMENTATION DU XYLOSE**
 (54) **Title: METABOLIC ENGINEERING OF XYLOSE FERMENTING EUKARYOTIC CELLS**

(57) **Abrégé/Abstract:**

The present invention relates to further genetic modifications in eukaryotic host cells that have been transformed to express a xylose isomerase that confers the host cell the ability of isomerising xylose to xylulose. The further genetic modifications are aimed at improving the efficiency of xylose metabolism and include e.g. reduction of unspecific aldose reductase activity, increased xylulose kinase activity and increased flux of the pentose phosphate pathway. The modified host cells of the invention are suitable for the production of a wide variety of fermentation products, including ethanol, in fermentation processes in which a source of xylose or a source of xylose and glucose are used as carbon source.

(72) **Inventeurs(suite)/Inventors(continued)**: DE LAAT, WILHELMUS THEODORUS ANTONIUS MARIA, NL;
VAN DIJKEN, JOHANNES PIETER, NL; PRONK, JACOBUS THOMAS, NL

(74) **Agent**: GOWLING LAFLEUR HENDERSON LLP

Abstract

The present invention relates to further genetic modifications in eukaryotic host cells that have been transformed to express a xylose isomerase that confers the host cell the ability of isomerising xylose to xylulose. The further genetic modifications are aimed at improving the efficiency of xylose metabolism and include e.g. reduction of unspecific aldose reductase activity, increased xylulose kinase activity and increased flux of the pentose phosphate pathway. The modified host cells of the invention are suitable for the production of a wide variety of fermentation products, including ethanol, in fermentation processes in which a source of xylose or a source of xylose and glucose are used as carbon source.

DEMANDES OU BREVETS VOLUMINEUX

**LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS
COMPREND PLUS D'UN TOME.**

CECI EST LE TOME _____ DE _____

NOTE: Pour les tomes additionels, veuillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

**THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE
THAN ONE VOLUME.**

THIS IS VOLUME 1 OF 2

NOTE: For additional volumes please contact the Canadian Patent Office.

Metabolic engineering of xylose-fermenting eukaryotic cellsField of the invention

5 The present invention relates to further genetic modifications in eukaryotic host cells that have been transformed to express a xylose isomerase that confers the host cell the ability of isomerising xylose to xylulose. The further genetic modifications are aimed at improving the efficiency of xylose metabolism and include e.g. reduction of
10 unspecific aldose reductase activity, increased xylulose kinase activity and increased flux of the pentose phosphate pathway. The modified host cells of the invention are suitable for the production of a wide variety of fermentation products in processes comprising xylose as carbon source.

Background of the invention

15 Economically viable ethanol production from the hemicellulose fraction of plant biomass requires the simultaneous conversion of both pentoses and hexoses at comparable rates and with high yields. Yeasts, in particular *Saccharomyces spp.*, are the most appropriate candidates for this process since they can grow fast on hexoses, both aerobically and anaerobically. Furthermore they are much more resistant to the
20 toxic environment of lignocellulose hydrolysates than (genetically modified) bacteria.

 In previous studies evidence has been provided that metabolic engineering of *S. cerevisiae* for xylose utilization, should be based on the introduction of xylose isomerase (XI, EC 5.3.1.5) Bruinenberg et al. (1983, Eur J. Appl. Microbiol. Biotechnol. 18: 287-292). In contrast to strains that are based on xylose reductase (XR, EC 1.1.1.21) and xylitol dehydrogenase (XD, EC 1.1.1.9), strains expressing XI
25 activity display high alcohol yields and hardly produce xylitol as has recently been demonstrated in WO 03/0624430 and Kuyper et al. (2004, FEMS Yeast Res. 4: 655-664). From a theoretical point of view this is not surprising since the route via XR and XD leads to an obstruction in the NADH balance that in the absence of oxygen, can be
30 relieved e.g. via xylitol formation.

 WO 03/0624430 discloses that the introduction of a functional *Piromyces* XI into *S. cerevisiae* allows slow metabolism of xylose via the endogenous xylulokinase (EC 2.7.1.17) encoded by *XKSI* and the enzymes of the non-oxidative part of the

pentose phosphate pathway and confers to the yeast transformants the ability to grow on xylose.

Kuyper et al. (*supra*) describe *S. cerevisiae* strains in which the *Piromyces XI* has been introduced and which are thereafter subjected to directed evolution in shake
5 flasks show improved rates of xylose fermentation, but still required oxygen for growth. Further selection via a regime of extreme oxygen limitation under xylose excess, followed by anaerobic selection resulted in a laboratory strain (RWB202-AFX) which fulfils at least one of the prerequisites for hemicellulose utilisation, namely an acceptable ethanol yield on xylose. However, the specific rate of ethanol production in
10 this strain is still unacceptably low. In particular, the specific sugar consumption rate during growth on xylose (345 mg xylose/g biomass/h) is still ten-fold lower than on glucose. Attempts to further improve strain RWB202-AFX via evolutionary engineering have failed so far.

WO 03/0624430 lists a number of alternative genetic modifications that may
15 result in further improvement of the specific rates of ethanol production and/or sugar consumption on xylose in host cells expressing the *Piromyces XI* gene to a level that would be required for commercial hemicellulose utilisation. These alternatives include:
(a) increase transport of xylose into the host cell; (b) increased xylulose kinase activity;
(c) increased flux of the pentose phosphate pathway; (d) decreased sensitivity to
20 catabolite repression; (e) increased tolerance to ethanol, osmolarity or organic acids; and, (f) reduced production of by-products (such as e.g. xylitol, glycerol and/or acetic acid). More specifically, WO 03/0624430 suggests to overexpress one or more of the genes encoding a hexose or pentose transporter, a xylulose kinase (such as the *S. cerevisiae XKSI*) an enzyme from the pentose phosphate pathway such as a
25 transaldolase (*TALI*) or a transketolase (*TKLI*) glycolytic enzymes, ethanologenic enzymes such as alcohol dehydrogenases, and/or to inactivate a hexose kinase gene, e.g. the *S. cerevisiae HXK2* gene, the *S. cerevisiae MIG1* or *MIG2* genes, the (unspecific) aldose reductase genes such as the *S. cerevisiae GRE3* gene, or genes for enzymes involved in glycerol metabolism such as the *S. cerevisiae* glycerol-phosphate
30 dehydrogenase 1 and/or 2 genes. WO 03/0624430 however does not disclose which of these many alternatives actually does produce an improvement in the specific rates of ethanol production and/or xylose consumption in host cells carrying the *Piromyces XI* gene.

Karhumaa et al. (2004, "Development of a Xylose-growing *Saccharomyces cerevisiae* strain expressing bacterial xylose isomerase", Poster presentation at the second meeting on Physiology of Yeasts and Filamentous Fungi; March 24-28 2004 Anglet, France. Page 43; and, 2004, "New Xylose-growing *Saccharomyces cerevisiae* strain for biofuel ethanol production", Oral presentation at the 26th Symposium on Biotechnology for fuels and chemicals, May 9-12, 2004 Chattanooga (TN), USA. Page 19) disclose a strain of *S. cerevisiae* expressing a bacterial XI from *Thermus thermophilus*. The strain further contains a number of the genetic modifications suggested in WO 03/0624430: overexpression of xylulose kinase and all four enzymes of the non-oxidative pentose phosphate pathway as well as inactivation of the *S. cerevisiae* unspecific aldose reductase gene (*GRE3*). However, despite these genetic modifications this strain is incapable of growth on xylose. Only after adaptation to aerobic growth on xylose a strain, TMB3050, was obtained that is capable of growth on xylose at a low rate ($\mu = 0.04 \text{ h}^{-1}$) and with a low specific xylose consumption rate of 4.3 mg xylose/g cells/h. Since undefined genetic modifications (accumulated during adaptation) are clearly required for growth on xylose in the first place, one cannot deduce from the work of Karhumaa et al., which, if any, of the defined genetic modifications (such as overexpression of xylulose kinase or any of the pentose phosphate pathway enzymes or inactivation of the aldose reductase gene) actually contribute to the ability of the adapted strain to grow on xylose.

It is therefore an object of the present invention to provide for eukaryotic host cells, such as fungal host cells, that are transformed with a XI gene that confers the ability to grow on xylose and which host cells have specific rates of xylose consumption and/or product (ethanol) formation that are compatible with commercial application of the host cells.

Description of the invention

Definitions

Xylose isomerase

The enzyme "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 vice versa. The enzyme is also known as a D-xylose ketoisomerase. Some xylose isomerases are also capable of catalysing the conversion between D-glucose and D-fructose and are

therefore sometimes referred to as glucose isomerase. Xylose isomerases require bivalent cations like magnesium or manganese as cofactor. Xylose isomerases of the invention may be further defined by their amino acid sequence as herein described below. Likewise xylose isomerases may be defined by the nucleotide sequences encoding the enzyme as well as by nucleotide sequences hybridising to a reference nucleotide sequence encoding a xylose isomerase as herein described below.

A unit (U) of xylose isomerase activity is herein 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).

10 Xylulose kinase

The enzyme "xylulose kinase" (EC 2.7.1.17) is herein defined as an enzyme that catalyses the reaction $ATP + D\text{-xylulose} = ADP + D\text{-xylulose 5-phosphate}$. 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 as herein described below. Likewise a xylulose kinase may be defined by the nucleotide sequences encoding the enzyme as well as by nucleotide sequences hybridising to a reference nucleotide sequence encoding a xylulose kinase as herein described below. A unit of xylulokinase activity is defined in Example 1.13 herein.

20 Ribulose 5-phosphate epimerase

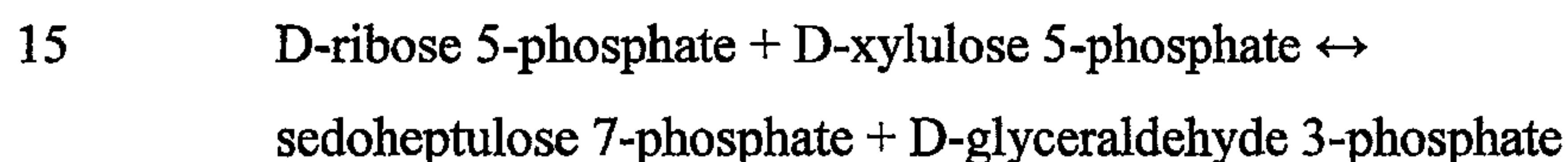
The enzyme "ribulose 5-phosphate epimerase" (5.1.3.1) is herein defined as an enzyme that catalyses the epimerisation of D-xylulose 5-phosphate into D-ribulose 5-phosphate and vice versa. The enzyme is also known as phosphoribulose epimerase; erythrose-4-phosphate isomerase; phosphoketopentose 3-epimerase;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

Ribulose 5-phosphate isomerase

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 phosphopentoisomerase; 5 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 of the invention may be further defined by its amino acid sequence as herein described below. Likewise a ribulose 5-phosphate isomerase may be defined by the nucleotide sequences encoding 10 the enzyme as well as by nucleotide sequences hybridising to a reference nucleotide sequence encoding a ribulose 5-phosphate isomerase as herein described below.

Transketolase

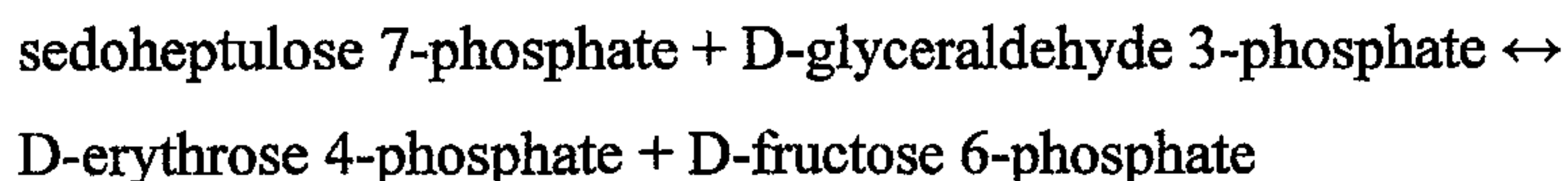
The enzyme "transketolase" (EC 2.2.1.1) is herein defined as an enzyme that catalyses the reaction:



and vice versa. The enzyme is also known as glycolaldehydetransferase or sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate glycolaldehydetransferase. A transketolase of the invention may be further defined by its amino acid sequence as 20 herein described below. Likewise a transketolase may be defined by the nucleotide sequences encoding the enzyme as well as by nucleotide sequences hybridising to a reference nucleotide sequence encoding a transketolase as herein described below.

Transaldolase

The enzyme "transaldolase" (EC 2.2.1.2) is herein defined as an enzyme that 25 catalyses the reaction:

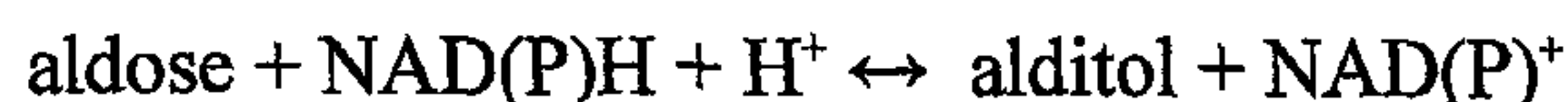


and vice versa. The enzyme is also known as dihydroxyacetone transferase; dihydroxyacetone synthase; formaldehyde transketolase; or sedoheptulose-7- 30 phosphate:D-glyceraldehyde-3-phosphate glyceronetransferase. A transaldolase of the invention may be further defined by its amino acid sequence as herein described below. Likewise a transaldolase may be defined by the nucleotide sequences encoding the

enzyme as well as by nucleotide sequences hybridising to a reference nucleotide sequence encoding a transaldolase as herein described below.

Aldose reductase

The enzyme "aldose reductase" (EC 1.1.1.21) is herein defined as any enzyme
5 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 or xylulose to xylitol. Unspecific aldose reductases catalyse the reaction:



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

20 Sequence identity and similarity

Sequence identity is herein defined as a relationship between two or more amino acid (polypeptide or protein) sequences or two or more nucleic acid (polynucleotide) sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between amino acid or nucleic acid sequences, as
25 the case may be, as determined by the match between strings of such sequences. "Similarity" between two amino acid sequences is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one polypeptide to the sequence of a second polypeptide. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in (Computational
30 Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular

Biology, von Heine, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM J. Applied Math., 48:1073 (1988).

Preferred methods to determine identity are designed to give the largest match
5 between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include e.g. the GCG program package (Devereux, J., et al., Nucleic Acids Research 12 (1):387 (1984)), BestFit, BLASTP, BLASTN, and FASTA (Altschul, S. F. et al., J. Mol. Biol. 215:403-410
10 (1990). The BLAST X program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990). The well-known Smith Waterman algorithm may also be used to determine identity.

Preferred parameters for polypeptide sequence comparison include the following:
15 Algorithm: Needleman and Wunsch, J. Mol. Biol. 48:443-453 (1970); Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, Proc. Natl. Acad. Sci. USA. 89:10915-10919 (1992); Gap Penalty: 12; and Gap Length Penalty: 4. A program useful with these parameters is publicly available as the "Ogap" program from Genetics Computer Group, located in Madison, WI. The aforementioned parameters are the
20 default parameters for amino acid comparisons (along with no penalty for end gaps).

Preferred parameters for nucleic acid comparison include the following:
Algorithm: Needleman and Wunsch, J. Mol. Biol. 48:443-453 (1970); Comparison matrix: matches=+10, mismatch=0; Gap Penalty: 50; Gap Length Penalty: 3. Available
25 as the Gap program from Genetics Computer Group, located in Madison, Wis. Given above are the default parameters for nucleic acid comparisons.

Optionally, in determining the degree of amino acid similarity, the skilled person may also take into account so-called "conservative" amino acid substitutions, as will be clear to the skilled person. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of
30 amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine,

tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulphur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-
5 valine, and asparagine-glutamine. Substitutional variants of the amino acid sequence disclosed herein are those in which at least one residue in the disclosed sequences has been removed and a different residue inserted in its place. Preferably, the amino acid change is conservative. Preferred conservative substitutions for each of the naturally occurring amino acids are as follows: Ala to ser; Arg to lys; Asn to gln or his; Asp to
10 glu; Cys to ser or ala; Gln to asn; Glu to asp; Gly to pro; His to asn or gln; Ile to leu or val; Leu to ile or val; Lys to arg; gln or glu; Met to leu or ile; Phe to met, leu or tyr; Ser to thr; Thr to ser; Trp to tyr; Tyr to trp or phe; and, Val to ile or leu.

Hybridising nucleic acid sequences

Nucleotide sequences encoding the enzymes of the invention may also be defined
15 by their capability to hybridise with the nucleotide sequences of SEQ ID NO.'s 9 – 16 and 18, respectively, under moderate, or preferably under stringent hybridisation conditions. Stringent hybridisation conditions are herein defined as conditions that allow a nucleic acid sequence of at least about 25, preferably about 50 nucleotides, 75 or 100 and most preferably of about 200 or more nucleotides, to hybridise at a
20 temperature of about 65°C in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength, and washing at 65°C in a solution comprising about 0.1 M salt, or less, preferably 0.2 x SSC or any other solution having a comparable ionic strength. Preferably, the hybridisation is performed overnight, i.e. at least for 10 hours and preferably washing is performed for at least one
25 hour with at least two changes of the washing solution. These conditions will usually allow the specific hybridisation of sequences having about 90% or more sequence identity.

Moderate conditions are herein defined as conditions that allow a nucleic acid sequences of at least 50 nucleotides, preferably of about 200 or more nucleotides, to
30 hybridise at a temperature of about 45°C in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength, and washing at room temperature in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength. Preferably, the

hybridisation is performed overnight, i.e. at least for 10 hours, and preferably washing is performed for at least one hour with at least two changes of the washing solution. These conditions will usually allow the specific hybridisation of sequences having up to 50% sequence identity. The person skilled in the art will be able to modify these
5 hybridisation conditions in order to specifically identify sequences varying in identity between 50% and 90%.

Operably linked

As used herein, the term "operably linked" refers to a linkage of polynucleotide elements in a functional relationship. A nucleic acid is "operably linked" when it is
10 placed into a functional relationship with another nucleic acid sequence. For instance, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the coding sequence. Operably linked means that the DNA sequences being linked are typically contiguous and, where necessary to join two protein coding regions, contiguous and in reading frame.

15 Promoter

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
20 polymerase, transcription initiation sites and any other DNA sequences, including, but not limited to transcription factor binding sites, repressor and activator protein binding sites, and any other sequences of nucleotides known to one of skill in the art to act directly or indirectly to regulate the amount of transcription from the promoter. A "constitutive" promoter is a promoter that is active under most environmental and
25 developmental conditions. An "inducible" promoter is a promoter that is active under environmental or developmental regulation.

Homologous

The term "homologous" when used to indicate the relation between a given (recombinant) nucleic acid or polypeptide molecule and a given host organism or host
30 cell, is understood to mean that in nature the nucleic acid or polypeptide molecule is produced by a host cell or organisms of the same species, preferably of the same variety or strain. If homologous to a host cell, a nucleic acid sequence encoding a polypeptide will typically be operably linked to another promoter sequence or, if

applicable, another secretory signal sequence and/or terminator sequence than in its natural environment. When used to indicate the relatedness of two nucleic acid sequences the term "homologous" means that one single-stranded nucleic acid sequence may hybridize to a complementary single-stranded nucleic acid sequence. The degree of hybridization may depend on a number of factors including the amount of identity between the sequences and the hybridization conditions such as temperature and salt concentration as discussed later. Preferably the region of identity is greater than about 5 bp, more preferably the region of identity is greater than 10 bp.

Heterologous

10 The term "heterologous" when used with respect to a nucleic acid (DNA or RNA) or protein refers to a nucleic acid or protein that does not occur naturally as part of the organism, cell, genome or DNA or RNA sequence in which it is present, or that is found in a cell or location or locations in the genome or DNA or RNA sequence that differ from that in which it is found in nature. Heterologous nucleic acids or proteins are not endogenous to the cell into which it is introduced, but has been obtained from another cell or synthetically or recombinantly produced. Generally, though not necessarily, such nucleic acids encode proteins that are not normally produced by the cell in which the DNA is transcribed or expressed. Similarly exogenous RNA encodes for proteins not normally expressed in the cell in which the exogenous RNA is present.

15 Heterologous nucleic acids and proteins may also be referred to as foreign nucleic acids or proteins. Any nucleic acid or protein that one of skill in the art would recognize as heterologous or foreign to the cell in which it is expressed is herein encompassed by the term heterologous nucleic acid or protein. The term heterologous also applies to non-natural combinations of nucleic acid or amino acid sequences, i.e. combinations where at least two of the combined sequences are foreign with respect to each other.

25 In this document and in its claims, the verb "to comprise" and its conjugations is used in its non-limiting sense to mean that items following the word are included, but items not specifically mentioned are not excluded. In addition, reference to an element by the indefinite article "a" or "an" does not exclude the possibility that more than one of the element is present, unless the context clearly requires that there be one and only one of the elements. The indefinite article "a" or "an" thus usually means "at least one".

30

Detailed description of the invention

The present invention relates to transformed eukaryotic host cells that have the ability of isomerising xylose to xylulose as e.g. described in WO 03/0624430. The ability of isomerising xylose to xylulose is conferred to the host cell by transformation
5 of the host cell with a nucleic acid construct comprising a nucleotide sequence encoding a xylose isomerase. The transformed host cell's ability to isomerise xylose into xylulose is the direct isomerisation of xylose to xylulose. This is understood to mean that xylose isomerised into xylulose in a single reaction catalysed by a xylose
10 isomerase, as opposed to the two step conversion of xylose into xylulose via a xylitol intermediate as catalysed by xylose reductase and xylitol dehydrogenase, respectively.

The nucleotide sequence encodes a xylose isomerase that is preferably expressed in active form in the transformed host cell. Thus, expression of the nucleotide sequence in the host cell produces a xylose isomerase with a specific activity of at least 10 U
15 xylose isomerase activity per mg protein at 30°C, preferably at least 20, 25, 30, 50, 100, 200, 300 or 500 U per mg at 30°C. The specific activity of the xylose isomerase expressed in the transformed host cell is herein defined as the amount of xylose isomerase activity units per mg protein of cell free lysate of the host cell, e.g. a yeast cell free lysate. Determination of the xylose isomerase activity, amount of protein and
20 preparation of the cell free lysate are as described in Example 1.13. Accordingly, expression of the nucleotide sequence encoding the xylose isomerase in the host cell produces a xylose isomerase with a specific activity of at least 50 U xylose isomerase activity per mg protein at 30°C, preferably at least 100, 200, 500, 750 or 1000 U per mg at 30°C.

25 Preferably, expression of the nucleotide sequence encoding the xylose isomerase in the host cell produces a xylose isomerase with a K_m for xylose that is less than 50, 40, 30 or 25 mM, more preferably, the K_m for xylose is about 20 mM or less.

A preferred nucleotide sequence encoding the xylose isomerase may be selected from the group consisting of:

- 30 (a) nucleotide sequences encoding a polypeptide comprising an amino acid sequence that has at least 60, 65, 70, 75, 80, 85, 90, 95, 97, 98, or 99% sequence identity with the amino acid sequence of SEQ ID NO. 1 and/or SEQ ID NO. 2;

- (b) nucleotide sequences comprising a nucleotide sequence that has at least 40, 50, 60, 70, 80, 90, 95, 97, 98, or 99% sequence identity with the nucleotide sequence of SEQ ID NO. 9 and/or SEQ ID NO. 10;
- (c) nucleotide sequences the complementary strand of which hybridises to a nucleic acid molecule sequence of (a) or (b);
- (d) nucleotide sequences the sequence of which differs from the sequence of a nucleic acid molecule of (c) due to the degeneracy of the genetic code.

The nucleotide sequence encoding the xylose isomerase may encode either a prokaryotic or an eukaryotic xylose isomerase, i.e. a xylose isomerase with an amino acid sequence that is identical to that of a xylose isomerase that naturally occurs in the prokaryotic or eukaryotic organism. The present inventors have found that the ability of a particular xylose isomerase to confer to a eukaryotic host cell the ability to isomerise xylose into xylulose does not depend so much on whether the isomerase is of prokaryotic or eukaryotic origin. Rather this depends on the relatedness of the isomerase's amino acid sequence to that of the *Piromyces* sequence (SEQ ID NO. 1). Surprisingly, the eukaryotic *Piromyces* isomerase is more related to prokaryotic isomerases than to other known eukaryotic isomerases. The *Piromyces* isomerase shares 61% amino acid identity with a *Xanthomonas* enzyme and 82% with a *Bacteroides* enzyme (SEQ ID NO. 2), whereas it only shares 49-52% identity with several plant xylose isomerases. No reports have issued of a plant xylose isomerase that is actively expressed in yeast. In contrast, in Example 3 herein we describe that a *Bacteroides* xylose isomerase confers to a eukaryotic host cell the ability to isomerise xylose into xylulose and to grow on xylose as sole carbon source. Therefore, a preferred nucleotide sequence encodes a xylose isomerase having an amino acid sequence that is related to the *Piromyces* sequence as defined above. A preferred nucleotide sequence encodes a fungal xylose isomerase (e.g. from a *Basidiomycete*), more preferably a xylose isomerase from an anaerobic fungus, e.g. a xylose isomerase from an anaerobic fungus that belongs to the families *Neocallimastix*, *Caecomyces*, *Piromyces*, *Orpinomyces*, or *Ruminomyces*. Alternatively, a preferred nucleotide sequence encodes a bacterial xylose isomerase, preferably a Gram-negative bacterium, more preferably an isomerase from the class *Bacteroides*, or from the genus *Bacteroides*, most preferably from *B. thetaiotaomicron* (SEQ ID NO. 2).

To increase the likelihood that the xylose isomerase is expressed in active form in a eukaryotic host cell such as yeast, the nucleotide sequence encoding the xylose isomerase may be adapted to optimise its codon usage to that of the eukaryotic host cell. The adaptiveness of a nucleotide sequence encoding the xylose isomerase (or other enzymes of the invention, see below) to the codon usage of the host cell may be expressed as codon adaptation index (CAI). The codon adaptation index is herein defined as a measurement of the relative adaptiveness of the codon usage of a gene towards the codon usage of highly expressed genes. The relative adaptiveness (w) of each codon is the ratio of the usage of each codon, to that of the most abundant codon for the same amino acid. The CAI index is defined as the geometric mean of these relative adaptiveness values. Non-synonymous codons and termination codons (dependent on genetic code) are excluded. CAI values range from 0 to 1, with higher values indicating a higher proportion of the most abundant codons (see Sharp and Li, 1987, *Nucleic Acids Research* 15: 1281-1295; also see: Jansen et al., 2003, *Nucleic Acids Res.* 31(8):2242-51). An adapted nucleotide sequence preferably has a CAI of at least 0.2, 0.3, 0.4, 0.5, 0.6 or 0.7.

A host cell for transformation with the nucleotide sequence encoding the xylose isomerase as described above, preferably is a host capable of active or passive xylose transport into the cell. The host cell preferably contains active glycolysis. The host cell may further contain an endogenous pentose phosphate pathway and may contain endogenous xylulose kinase activity so that xylulose isomerised from xylose may be metabolised to pyruvate. The host further preferably contains enzymes for conversion of pyruvate to a desired fermentation product such as ethanol, lactic acid, 3-hydroxypropionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, β -lactam antibiotics and cephalosporins.

A preferred host cell is a host cell that is naturally capable of alcoholic fermentation, preferably, anaerobic alcoholic fermentation. The host cell further preferably has a high tolerance to ethanol, a high tolerance to low pH (i.e. capable of growth at a pH lower than 5, 4, 3, or 2,5) and towards organic acids like lactic acid, acetic acid or formic acid and sugar degradation products such as furfural and hydroxymethylfurfural, and a high tolerance to elevated temperatures. Any of these characteristics or activities of the host cell may be naturally present in the host cell or may be introduced or modified by genetic modification. A suitable host cell is a

eukaryotic microorganism like e.g. a fungus, however, most suitable as host cell are yeasts or filamentous fungi.

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. Preferred yeasts as host cells belong to the genera *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces*, and *Yarrowia*. Preferably the yeast is capable of anaerobic fermentation, more preferably 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 present invention are morphologically, physiologically, and genetically distinct from yeasts. Vegetative growth by filamentous fungi is by hyphal elongation and carbon catabolism of most filamentous fungi is obligately aerobic. Preferred filamentous fungi as host cells belong to the genera *Aspergillus*, *Trichoderma*, *Humicola*, *Acremonium*, *Fusarium*, and *Penicillium*.

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 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*, *K. fragilis*.

The host cell of the invention is thus a host cell that is transformed with a nucleic acid construct comprising the nucleotide sequence encoding the xylose isomerase as defined above. The nucleic acid construct comprising the xylose isomerase coding sequence preferably is capable of expression of the xylose isomerase in the host cell. To this end the nucleic acid construct may be constructed as described in e.g. WO 03/0624430. The host cell may comprise a single but preferably comprises multiple copies of the nucleic acid construct. The nucleic acid construct may be maintained

episomally and thus comprise a sequence for autonomous replication, such as an ARS sequence. Suitable episomal nucleic acid constructs may e.g. be based on the yeast 2μ or pKD1 (Fleer et al., 1991, *Biotechnology* 9:968-975) plasmids. Preferably, however, the nucleic acid construct is integrated in one or more copies into the genome of the host cell. Integration into the host cell's genome may occur at random by illegitimate recombination but preferably nucleic acid construct is integrated into the host cell's genome by homologous recombination as is well known in the art of fungal molecular genetics (see e.g. WO 90/14423, EP-A-0 481 008, EP-A-0 635 574 and US 6,265,186).

In a first aspect of the invention, the host cell of the invention comprises a genetic modification that increases the flux of the pentose phosphate pathway. In particular, the genetic modification causes an increased flux of the non-oxidative part 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 1.1, 1.2, 1.5, 2, 5, 10 or 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 (μ_{\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.

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
5 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-
10 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
15 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,
20 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
25 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 we have found that host cells overexpressing only the transketolase and the transaldolase already have the same anaerobic growth rate on xylose as do host cells
30 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.

There are various means available in the art for overexpression of enzymes in the host cells of the invention. In particular, an enzyme may be overexpressed by increasing the copynumber 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, preferably under conditions where xylose or xylose and glucose are available as carbon sources, more preferably as major carbon sources (i.e. more than 50% of the available carbon source consists of xylose or xylose and glucose), most preferably as sole carbon sources. Suitable promoters in this context include both constitutive and inducible natural promoters as well as engineered promoters. A preferred promoter for use in the present invention will in addition be insensitive to catabolite (glucose) repression and/or will preferably not require xylose for induction.

Promoters having these characteristics 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 (PPK), triose phosphate isomerase (TPI), glyceraldehyde-3-phosphate dehydrogenase (GPD, TDH3 or GAPDH), pyruvate kinase (PYK), 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 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.

The coding sequence used for overexpression of the enzymes preferably is
5 homologous to the host cell of the invention. However, coding sequences that are heterologous to the host cell of the invention may likewise be applied.

A nucleotide sequence used for overexpression of ribulose-5-phosphate isomerase in the host cell of the invention is a nucleotide sequence encoding a polypeptide with ribulose-5-phosphate isomerase activity, whereby preferably the
10 polypeptide has an amino acid sequence having at least 50, 60, 70, 80, 90 or 95% identity with SEQ ID NO. 4 or whereby the nucleotide sequence is capable of hybridising with the nucleotide sequence of SEQ ID NO. 12, under moderate conditions, preferably under stringent conditions.

A nucleotide sequence used for overexpression of ribulose-5-phosphate
15 epimerase in the host cell of the invention is a nucleotide sequence encoding a polypeptide with ribulose-5-phosphate epimerase activity, whereby preferably the polypeptide has an amino acid sequence having at least 50, 60, 70, 80, 90 or 95% identity with SEQ ID NO. 5 or whereby the nucleotide sequence is capable of hybridising with the nucleotide sequence of SEQ ID NO. 13, under moderate
20 conditions, preferably under stringent conditions.

A nucleotide sequence used for overexpression of transketolase in the host cell of the invention is a nucleotide sequence encoding a polypeptide with transketolase activity, whereby preferably the polypeptide has an amino acid sequence having at least 50, 60, 70, 80, 90 or 95% identity with SEQ ID NO. 6 or whereby the nucleotide
25 sequence is capable of hybridising with the nucleotide sequence of SEQ ID NO. 14, under moderate conditions, preferably under stringent conditions.

A nucleotide sequence used for overexpression of transaldolase in the host cell of the invention is a nucleotide sequence encoding a polypeptide with transaldolase activity, whereby preferably the polypeptide has an amino acid sequence having at least
30 50, 60, 70, 80, 90 or 95% identity with SEQ ID NO. 7 or whereby the nucleotide sequence is capable of hybridising with the nucleotide sequence of SEQ ID NO. 15, under moderate conditions, preferably under stringent conditions.

Overexpression of an enzyme, when referring to the production of the enzyme in a genetically modified host 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. Overexpression of an enzyme is thus preferably determined by measuring the level of the enzyme's specific activity in the host cell using appropriate enzyme assays as described herein. Alternatively, overexpression of the enzyme may determined indirectly by quantifying the specific steady state level of enzyme protein, e.g. using antibodies specific for the enzyme, or by quantifying the specific steady level of the mRNA coding for the enzyme. The latter may particularly be suitable for enzymes of the pentose phosphate pathway for which enzymatic assays are not easily feasible as substrates for the enzymes are not commercially available. Preferably in the host cells of the invention, an enzyme to be overexpressed is overexpressed by at least a factor 1.1, 1.2, 1.5, 2, 5, 10 or 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.

In a second aspect of the invention, the host cell of the invention comprises a genetic modification that increases the specific xylulose kinase activity. Preferably the genetic modification 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 of the invention is a nucleotide sequence encoding a polypeptide with xylulose kinase activity, whereby preferably the polypeptide has an amino acid sequence having at least 50, 60, 70, 80, 90 or 95% identity with SEQ ID NO. 3 or whereby the nucleotide sequence is capable of hybridising with the nucleotide

sequence of SEQ ID NO. 11, under moderate conditions, preferably under stringent conditions.

A particularly preferred xylulose kinase is a xylose kinase that is related to the xylulose kinase from *Piromyces* (*xylB*; see WO 03/0624430). This *Piromyces* xylulose
5 kinase is actually more related to prokaryotic kinase than to all of the known eukaryotic kinases such as the yeast kinase (SEQ ID NO. 3). The eukaryotic xylulose kinases have been indicated as non-specific sugar kinases, which have a broad substrate range that includes xylulose. In contrast, the prokaryotic xylulose kinases, to which the *Piromyces* kinase is most closely related, have been indicated to be more specific kinases for
10 xylulose, i.e. having a narrower substrate range. Therefore, a more preferred nucleotide sequence for use in overexpression of xylulose kinase in the host cell of the invention is a nucleotide sequence encoding a polypeptide with xylulose kinase activity, whereby preferably the polypeptide has an amino acid sequence having at least 45, 50, 55, 60, 65, 70, 80, 90 or 95% identity with SEQ ID NO. 17 or whereby the nucleotide
15 sequence is capable of hybridising with the nucleotide sequence of SEQ ID NO. 18, under moderate conditions, preferably under stringent conditions.

In the host cells of the invention, genetic modification that increases the specific xylulose kinase activity may be combined with any of the modifications increasing the flux of the pentose phosphate pathway as described above, but this combination is not
20 essential for the invention. Thus, a host cell of the invention comprising only a genetic modification that increases the specific xylulose kinase activity is specifically included in the invention. The various means available in the art for 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
25 cells of the invention, a xylulose kinase to be overexpressed is overexpressed by at least a factor 1.1, 1.2, 1.5, 2, 5, 10 or 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
30 state level of the transcript coding for the enzyme.

In a third aspect of the invention, the host cell of the invention comprises a genetic modification that reduces unspecific aldose reductase activity in the host cell. Preferably, unspecific aldose reductase activity is 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 modifications reduce or inactivate the expression of each endogenous copy of a gene encoding an unspecific aldose reductase in the host cell. Host cells may comprise multiple copies of genes encoding unspecific 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. A nucleotide sequence encoding an aldose reductase whose activity is to be reduced in the host cell of the invention is a nucleotide sequence encoding a polypeptide with aldose reductase activity, whereby preferably the polypeptide has an amino acid sequence having at least 50, 60, 70, 80, 90 or 95% identity with SEQ ID NO. 8 or whereby the nucleotide sequence is capable of hybridising with the nucleotide sequence of SEQ ID NO. 16 under moderate conditions, preferably under stringent conditions.

In the host cells of the invention, genetic modification that reduces unspecific aldose reductase activity in the host cell may be combined with any of the modifications increasing the flux of the pentose phosphate pathway and/or with any of the modifications increasing the specific xylulose kinase activity in the host cells as described above, but these combinations are not essential for the invention. Thus, a host cell of the invention comprising only a genetic modification that reduces unspecific aldose reductase activity in the host cell is specifically included in the invention.

In a further aspect the invention relates to modified host cells that are further adapted to xylose utilisation by selection of mutants, either spontaneous or induced (e.g. by radiation or chemicals), for growth on xylose, preferably on xylose as sole carbon source, and more preferably under anaerobic conditions. Selection of mutants may be performed by serial passaging 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 as is described in Example 4 herein.

In a preferred host cell of the invention 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 xylose as carbon source, preferably as sole carbon source, and preferably under anaerobic conditions. Preferably the modified host cell
5 produce essentially no xylitol, e.g. the xylitol produced is below the detection limit or e.g. less than 5, 2, 1, 0.5, or 0.3 % of the carbon consumed on a molar basis.

Preferably the modified host cell has the ability to grow on xylose as sole carbon source at a rate of at least 0.05, 0.1, 0.2, 0,25 or 0,3 h⁻¹ under aerobic conditions, or, if applicable, at a rate of at least 0.03, 0.05, 0.07, 0.08, 0.09, 0.1, 0.12, 0.15 or 0.2 h⁻¹
10 under anaerobic conditions. Preferably the modified host cell has the ability to grow on a mixture of glucose and xylose (in a 1:1 weight ratio) as sole carbon source at a rate of at least 0.05, 0.1, 0.2, 0,25 or 0,3 h⁻¹ under aerobic conditions, or, if applicable, at a rate of at least 0.03, 0.05, 0.1, 0.12, 0.15, or 0.2 h⁻¹ under anaerobic conditions.

Preferably, the modified host cell has a specific xylose consumption rate of at
15 least 346, 350, 400, 500, 600, 750, or 1000 mg xylose/g cells/h. Preferably, the modified host cell has a yield of fermentation product (such as ethanol) on xylose that is at least 55, 60, 70, 80, 85, 90, 95 or 98% of the host cell's yield of fermentation product (such as ethanol) on glucose. More preferably, the modified host cell's yield of fermentation product (such as ethanol) on xylose is equal to the host cell's yield of
20 fermentation product (such as ethanol) on glucose. Likewise, the modified host cell's biomass yield on xylose is preferably at least 55, 60, 70, 80, 85, 90, 95 or 98% of the host cell's biomass yield on glucose. More preferably, the modified host cell's biomass yield on xylose is equal to the host cell's biomass yield on glucose. It is understood that in the comparison of yields on glucose and xylose both yields are compared under
25 aerobic conditions or both under anaerobic conditions.

In a preferred aspect, the modified host cell of the invention is a host cell for the production of ethanol. In another aspect the invention relates to a transformed host cell for the production of fermentation products other than ethanol. Such non-ethanolic fermentation products include in principle any bulk or fine chemical that is producible
30 by a eukaryotic microorganism such as a yeast or a filamentous fungus. Such fermentation products include e.g. lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, β -lactam antibiotics and cephalosporins. A preferred modified host cell of the invention

for production of non-ethanolic fermentation products is a host cell that contains a genetic modification that results in decreased alcohol dehydrogenase activity.

In a further aspect the invention relates to fermentation processes in which the modified host cells of the invention are used for the fermentation of a carbon source comprising a source of xylose, such as xylose. In addition to a source of xylose the carbon source in the fermentation medium may also comprise a source of glucose. The source of xylose or glucose may be xylose or glucose as such or may be any carbohydrate oligo- or polymer comprising xylose or glucose units, such as e.g. lignocellulose, xylans, cellulose, starch and the like. For release of xylose or glucose units from such carbohydrates, appropriate carbohydrases (such as xylanases, glucanases, amylases and the like) may be added to the fermentation medium or may be produced by the modified host cell. In the latter case the modified host cell may be genetically engineered to produce and excrete such carbohydrases. An additional advantage of using oligo- or polymeric sources of glucose is that it enables to maintain a low(er) concentration of free glucose during the fermentation, e.g. by using rate-limiting amounts of the carbohydrases. This, in turn, will prevent repression of systems required for metabolism and transport of non-glucose sugars such as xylose. In a preferred process the modified host cell ferments both the xylose and glucose, preferably simultaneously in which case preferably a modified host 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 modified host cell. Compositions of fermentation media for growth of microorganisms such as yeasts are well known in the art.

The fermentation process is a process for the production of a fermentation product such as e.g. ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, β -lactam antibiotics, such as Penicillin G or Penicillin V and fermentative derivatives thereof, and cephalosporins. 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 5, 2.5 or 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
5 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, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, β -lactam antibiotics and cephalosporins.

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

A preferred process is a process for the production of ethanol, whereby the process comprises the steps of: (a) fermenting a medium containing a source of xylose with a modified host cell as defined above, whereby the host cell ferments xylose to ethanol; and optionally, (b) recovery of the ethanol. The fermentation medium may also
20 comprise a source of glucose that is also fermented to ethanol. In the process the volumetric ethanol productivity is preferably at least 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 5.0 or 10.0 g ethanol per litre per hour. The ethanol yield on xylose and/or glucose in the process preferably is at least 50, 60, 70, 80, 90, 95 or 98%. The ethanol yield is herein defined as a percentage of the theoretical maximum yield, which, for glucose and
25 xylose is 0.51 g. ethanol per g. glucose or xylose.

In a further aspect the invention relates to a process for producing a fermentation product selected from the group consisting of lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, β -lactam antibiotics and cephalosporins. The process preferably
30 comprises the steps of (a) fermenting a medium containing a source of xylose with a modified host cell as defined herein above, whereby the host cell ferments xylose to the fermentation product, and optionally, (b) recovery of the fermentation product. In a preferred process, the medium also contains a source of glucose.

Genetic modifications

For overexpression of enzymes in the host cells of the inventions as described above, as well as for additional genetic modification of host cells, preferably yeasts, host cells are transformed with the various nucleic acid constructs of the invention by methods well known in the art. Such methods are e.g. 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 fungal host cells are known from e.g. EP-A-0 635 574, WO 98/46772, WO 99/60102 and WO 00/37671.

Promoters for use in the nucleic acid constructs for overexpression of enzymes in the host cells of the invention have been described above. In the nucleic acid constructs for overexpression, the 3'-end of the nucleotide acid sequence encoding the enzyme(s) preferably is operably linked to a transcription terminator sequence. Preferably 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. The transcription termination sequence further preferably comprises a polyadenylation signal.

Optionally, a selectable marker may be present in the nucleic acid construct. 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 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). Although the of antibiotic resistance markers may be most convenient for the transformation of polyploid host cells, preferably however, non-antibiotic resistance markers are used, such as auxotrophic markers (URA3, TRP1, 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-0 635 574 and are based on the use of bidirectional markers such as the *A. nidulans* amdS (acetamidase) gene or the yeast URA3 and LYS2 genes. Alternatively, a screenable marker such as Green Fluorescent Protein, *lacZ*, 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 of the invention include, but are not limited to, one or more leader 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. Suitable episomal nucleic acid constructs may e.g. be based on the yeast 2 μ or pKD1 (Fleer et al., 1991, *Biotechnology* 9:968-975) plasmids. Alternatively the nucleic acid construct may comprise sequences for integration, preferably by homologous recombination. Such sequences may thus be sequences homologous to the target site for integration in the host cell's genome. The nucleic acid constructs of the invention can be provided in a manner known per se, which generally involves techniques such as restricting and linking nucleic acids/nucleic acid sequences, for which reference is made to the 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 inactivation and gene disruption in yeast or filamentous fungi are well known in the art (see e.g. Fincham, 1989, *Microbiol Rev.* 53(1):148-70 and EP-A-0 635 574).

Description of the figures

Figure 1. Typical graph of anaerobic growth of strain RWB 212 in fermenters on synthetic medium with 2% (w/v) xylose as the carbon source, duplicate experiments differed by less than 5%. Panel A: Xylose (●), ethanol (○), glycerol (■) and cumulative CO₂ produced per litre as deduced from gas analysis (-). Panel B: dry weight (●), acetate (○), xylitol (□), succinate (▲), lactate (△).

Figure 2. Typical graph of anaerobic growth of strain RWB 212 in fermenters on synthetic medium with 2% (w/v) glucose and 2% (w/v) xylose as the carbon source, duplicate experiments differed by less than 5%. Panel A: Glucose (●), xylose (○), ethanol (■), glycerol (□) and cumulative CO₂ produced per litre as deduced from gas analysis (–). Panel B: dry weight (●), acetate (○), xylitol (■), lactate (□) succinate (▲).

Figure 3. Panel A: Residual xylose concentrations during anaerobic chemostat cultivation of RWB 212 on 30 g/l xylose as the carbon source. Data presented are the average of two independent chemostats and the experimental deviations.

10 Panel B: Culture dry weight during anaerobic chemostat cultivation of RWB 212 on 30 g/l xylose as the carbon source. Data presented are the average of two independent chemostats and the experimental deviations.

Figure 4. Carbon dioxide production profiles of three xylose metabolising strains in anaerobic batch cultivations on glucose and xylose (20 g/l each). Exact experimental conditions varied so actual numeric values may not be compared.

15

Figure 5. Concentrations of glucose, xylose, ethanol, glycerol and CO₂ measured during two independent anaerobic fermentor batches on 2% glucose and 2% xylose of selected strains originating from RWB 212.

Examples1. Materials and methods1.1. Plasmid construction

In order to integrate the *TPII* promoter in front of our target genes several
 5 plasmids were constructed. First the *TPII* promoter was cut as a *XhoI-EcoRV* fragment
 from pYX012-Aat (A.A. Winkler, derivative of pYX012 (R&D Systems, Minneapolis,
 MN, USA)) and ligated to pUG6 [3] cut with *Sall-PvuII*. This gave us pUG6P_{TPI},
 which could then be used to PCR a Kanlox-P_{TPI} integration cassette.

In cases where putative ORF's were located very close to the ATG of the target
 10 genes, we cloned those genes into pUG6P_{TPI}. A 0.8 kb *RPE1* fragment and a 2.3 kb
TKL1 fragment were isolated from gel and cut with *EcoRI* and *XhoI* (present in the
 primers, see Table 3) and ligated into pUG6P_{TPI} digested with *EcoRI-Sall*, resulting in
 pUG6P_{TPI}-RPE1 and pUG6P_{TPI}-TKL1.

In order to increase the activity of xylulokinase the *XKS1* gene was amplified by
 15 PCR as a *SpeI-Sall* fragment (sites in the primers, see Table 3) and cloned into
 p415ADH [4] cut with *XbaI-XhoI*, resulting in p415ADHXKS.

Restriction endonucleases (New England Biolabs, Beverly, MA, USA and Roche,
 Basel, Switzerland) and DNA ligase (Roche) were used according to the manufacturers'
 specifications. Plasmid isolation from *E. coli* was performed with the Qiaprep spin
 20 miniprep kit (Qiagen, Hilden, Germany). DNA fragments were separated on a 1%
 agarose (Sigma, St. Louis, MO, USA) gel in 1×TBE [5]. Isolation of fragments from
 gel was carried out with the Qiaquick gel extraction kit (Quiagen). Amplification of
RPE1, *TKL1* and *XKS1* was done with Vent_R DNA polymerase (New England Biolabs)
 according to the manufacturer's specification. The template was chromosomal DNA of
 25 CEN.PK113-7D (wt). The polymerase chain reaction (PCR) was performed in a
 Biometra TGradient Thermocycler (Biometra, Göttingen, Germany) with the following
 settings: 30 cycles of 1 min annealing at 60°C, 3 min extension at 75°C and 1 min
 denaturing at 94°C.

1.2. Strains and media

30 The *Saccharomyces cerevisiae* strain used in this study is RWB212 (*MATA ura3-52 leu2-112 loxP-P_{TPI}::(-266,-1)TAL1 gre3::hphMX pUGP_{TPI}-TKL1 pUGP_{TPI}-RPE1 KanloxP-P_{TPI}::(-?, -1)RKII*), which is derived from CEN.PK102-3A (*MATA ura3-52 leu2-112*).

During construction strains were maintained on complex (YP: 10 g l⁻¹ yeast extract (BD Difco), 20 g l⁻¹ peptone (BD Difco)) or synthetic medium (MY) [6] supplemented with glucose (2%) as carbon source (YPD or MYD) and 1.5% agar in the case of plates. After transformation integrants were selected by plating on YPD containing geneticin (G418) (Invitrogen/GIBCO) at 200 µg/ml or hygromycin B (Roche Diagnostics GmbH, Mannheim, Germany) at 300 µg/ml. After transformation with plasmids strains were plated on MYD. Transformations of yeast were done according to Gietz and Woods [7].

Plasmids were amplified in *Escherichia coli* strain XL-1 blue (Stratagene, La Jolla, CA, USA). Transformation was performed according to Inoue et al. [8]. *E. coli* was grown on LB (Luria-Bertani) plates or in liquid TB (Terrific Broth) medium for the isolation of plasmids [5].

1.3. Strain construction

For *TAL1* and *RKII* integration of the *TPII* promoter 5' of the open reading frame was done by amplifying a PCR fragment with the KanMX marker and the *TPII* promoter and directing it to the target location via homologous ends. The PCR was performed with *Taq* DNA polymerase (Amersham Pharmacia, Piscataway, USA) according to the manufacturer's specifications. The template was pUG6P_{TPI} with P_{TALdisA} and P_{TALdisB} or P_{RKIdisA} and P_{RKIdisB} (Table 3) as primers.

In the case of *TKL1* and *RPE1*, plasmids pUG6P_{TPI}-*TKL1* and pUG6P_{TPI}-*RPE1* were linearized with *PvuII* and *SalI* respectively and integrated into the genome. Correct integration of the constructs was verified colony PCR with *TAL1* intern + KanA for *TAL1* and P_{TPI} primer + "intern" for *TKL1*, *RPE1*, and *RPII*. The "intern" primers anneal downstream of the integrated constructs, while P_{TPI} primer anneals at the 3' end of the *TPII* promoter.

After integration of a construct the KanMX marker was removed with the cre recombinase. To this end strains were transformed with pSH47 [3]. Colonies with the plasmid were resuspended in YP with 1% galactose and incubated for 1 hour at 30°C. Then about 200 cells were plated on YPD. The resulting colonies were checked for loss of the KanMX marker and pSH47 (*URA3*).

In addition the *GRE3* gene was replaced by the hphMX cassette from pAG32, conferring hygromycin resistance [9]. The hphMX cassette was amplified using oligo's 5'gre3::Kanlox and 3'gre3::Kanlox. Correct integration was verified using by PCR

with 5'GRE3 + KanA and 3'GRE3 + KanB (Table 3). KanA and KanB anneal to the *A. gossipi* TEF promoter and terminator respectively, while the other primers anneal outside of the *GRE3* open reading frame.

Colony PCR was done with *Taq* DNA polymerase (Amersham Pharmacia, Piscataway, USA) according to the manufacturer's specifications. As template cells were resuspended in 2.5 µl 0.02M NaOH to which the PCR reaction mixture was added. The PCR was performed in a Biometra TGradient Thermocycler (Biometra, Göttingen, Germany) with the following settings: 30 cycles of 1 min annealing at 60°C, 3 min extension at 72°C and 1 min denaturing at 94°C.

The resulting strain, RWB212 (*MATA ura3-52 leu2-112 loxP-P_{TPI}::(-266,-1)TAL1 gre3::hphMX pUGP_{TPI}-TKL1 pUGP_{TPI}-RPE1 KanloxP-P_{TPI}::(-?, -1)RKII*), was then transformed with pAKX002, a multicopy vector containing the *Piromyces* sp. E2 *XylA* behind the *TPII* promoter, as well as p415ADHXKS. Which gave us RWB217 (*MATA ura3-52 leu2-112 loxP-P_{TPI}::(-266,-1)TAL1 gre3::hphMX pUGP_{TPI}-TKL1 pUGP_{TPI}-RPE1 KanloxP-P_{TPI}::(-?, -1)RKII* {pAKX002, p415ADHXKS}).

1.4. Strain maintenance

Stock cultures were grown at 30°C in shake flasks on synthetic medium [6] supplemented with 20 g of glucose l⁻¹. When stationary phase was reached, sterile glycerol was added to 30% (vol/vol), and 2-ml aliquots were stored in sterile vials at -80°C.

1.5. Cultivation and media

Shake-flask cultivation was performed at 30°C in a synthetic medium [6]. The pH of the medium was adjusted to 6.0 with 2 M KOH prior to sterilization. Precultures were prepared by inoculating 100 ml medium containing 20 g l⁻¹ xylose in a 500 ml shake-flask with a frozen stock culture. After 24 to 48 h incubation at 30°C in an orbital shaker (200 rpm), this culture was used to inoculate either shake-flask cultures or fermenter cultures. The synthetic medium for anaerobic cultivation was supplemented with 0.01 g l⁻¹ ergosterol and 0.42 g l⁻¹ Tween 80 dissolved in ethanol [10,11], this resulted in 11-13 mM ethanol in the medium.

1.6. Anaerobic batch cultivation in fermenters

Anaerobic batch cultures were carried out in 2-litre laboratory fermenters (Applikon, Schiedam, The Netherlands) equipped with Norprene tubing, with a working volume of 1.5 litres, at 30°C and at pH 5.0. Cultures were stirred at 800 rpm

and sparged with 0.5 l min^{-1} of high-grade nitrogen (<5 ppm oxygen). The synthetic medium was supplemented with the anaerobic growth factors ergosterol and Tween 80 (0.01 and 0.42 g l^{-1} , respectively) as well as $100 \mu\text{l l}^{-1}$ of silicone antifoam (BDH, Poole, UK).

5 1.7. Determination of culture dry weight

Culture samples (10.0 ml) were filtered over preweighed nitrocellulose filters (pore size $0.45 \mu\text{m}$; Gelman laboratory, Ann Arbor, USA). After removal of medium the filters were washed with demineralised water and dried in a microwave oven (Bosch, Stuttgart, Germany) for 20 min at 360 W and weighed. Duplicate
10 determinations varied by less than 1 %.

1.8. Gas analysis

Exhaust gas was cooled in a condensor (2°C) and dried with a Permapure dryer type MD-110-48P-4 (Permapure, Toms River, USA). O_2 and CO_2 concentrations were determined with a NGA 2000 analyser (Rosemount Analytical, Orrville, USA).
15 Exhaust gas-flow rate and specific oxygen-consumption and carbon-dioxide production rates were determined as described previously [12,13]. In calculating these biomass-specific rates, a correction was made for volume changes caused by withdrawing culture samples.

1.9. Metabolite analysis

20 Glucose, xylose, xylitol, organic acids, glycerol and ethanol were detected by HPLC analysis on a Waters Alliance 2690 HPLC (Waters, Milford, USA) containing a Biorad HPX 87H column (Biorad, Hercules, USA). The column was eluted at 60°C with $0.5 \text{ g l}^{-1} \text{H}_2\text{SO}_4$ at a flow rate of 0.6 ml min^{-1} . Detection was by means of a Waters 2410 refractive-index detector and a Waters 2487 UV detector. Xylulose was
25 determined enzymatically in the following manner. The reaction mixture consisted of 100 mM Tris-HCl buffer (pH7.5) with 10 mM MgCl_2 , 0.30 mM NADH and an adequate amount of sample (1 ml total volume) the assay was started by the addition of 0.2 U sorbitol dehydrogenase (Sigma, St Louis, USA). The xylulose concentration was calculated using an absorption coefficient of $6.3 \text{ mM}^{-1} \text{ cm}^{-1}$ for NADH.

30 1.10. Carbon recoveries and ethanol evaporation

Carbon recoveries were calculated as carbon in products formed divided by the total amount of sugar carbon consumed and were based on a carbon content of biomass of 48%. To correct for ethanol evaporation during the fermentations, the amount of

ethanol produced was assumed to be equal to the measured cumulative production of CO₂ minus the CO₂ production that occurred due to biomass synthesis (5.85 mmol CO₂ per gram biomass [14]) and the CO₂ associated with acetate formation as described previously [2].

5 1.11 Microarray analysis

Sampling of cells from chemostats, probe preparation and hybridization to Affymetrix Genechip® microarrays were performed as described previously [15]. The results for each growth condition were derived from three independently cultured replicates.

10 1.12. Data acquisition and analysis

Acquisition and quantification of array images and data filtering were performed using the Affymetrix software packages: Microarray Suite v5.0, MicroDB v3.0 and Data Mining Tool v3.0.

15 Before comparison, all arrays were globally scaled to a target value of 150 using the average signal from all gene features using Microarray Suite v5.0. From the 9,335 transcript features on the YG-S98 arrays a filter was applied to extract 6,383 yeast open reading frames of which there were 6,084 different genes. This discrepancy was due to several genes being represented more than once when sub-optimal probe sets were used in the array design.

20 To represent the variation in triplicate measurements, the coefficient of variation (C.V.; standard deviation divided by the mean) was calculated as previously described by Boer et al. [16].

For further statistical analyses Microsoft Excel running the Significant Analysis of Microarrays (SAM v1.12) add in was used [17] for possible pair wise comparisons
25 of the eight data sets. Genes were considered as being changed in expression if they were called significantly changed using SAM (expected median false-discovery rate (FDR) of 1%) by at least 2-fold from each other condition. Hierarchical clustering of the obtained sets of significantly changed expression levels was subsequently performed by Genespring (Silicon Genetics).

30 1.13 Enzyme assays

Xylose isomerase activity was assayed at 37°C in a reaction mixture containing 50 mM phosphate buffer (pH 7.0), 10 mM xylose, 10 mM MgCl₂ and a suitable amount of cell-free extract. The amount of xylulose formed was determined by the cysteine-

carbazole method (9). Alternatively xylose isomerase activity was assayed at 30°C enzyme assay as developed by Kersters-Hildersson et al. (Kinetic characterization of D-xylose isomerases by enzymatic assays using D-sorbitol dehydrogenase. *Enz. Microb. Technol.* 9 (1987) 145-148). The in vitro activity of xylose isomerase in the cell-free extracts of transformed *S.cerevisiae* strains is dependent on bivalent cations (Mg²⁺ or Co²⁺).

Xylulose kinase and xylose reductase activities were assayed as described by Witteveen et al. (28). One unit of activity is defined as the amount of enzyme producing 1 nmol of xylulose per min under the assay conditions. Xylulose formed was determined by the method of Dische and Borenfreund (Dische and Borenfreund, 1951, *J. Biol. Chem.* 192: 583-587) or by HPLC using a Biorad HPX-87N Column operated at 80°C and eluated at 0.6 ml/min using 0.01 M Na₂HPO₄ as the eluens. Xylose and xylulose were detected by a Refractive Index detector at an internal temperature of 60°C.

Specific activity is expressed as units per mg protein. Protein was determined with the Bio-Rad protein reagent (Bio-Rad Laboratories, Richmond, CA, USA) with bovine γ -globulin as a standard.

2. Results

2.1 Overexpression of the pentose phosphate pathway (PPP) genes

Previously we have shown that expressing a fungal xylose isomerase in *Saccharomyces cerevisiae* is in principle enough to allow anaerobic growth of this yeast on xylose as the sole carbon source provided that sufficient selective pressure is applied [2]. The selected strain still however, did not perform up to industrial requirements (Table 1).

In order to investigate the possibility of rate limiting steps in pentose phosphate metabolism it was decided to construct a strain overproducing all the enzymes required to convert xylose into fructose-6-phosphate and glyceraldehyde-3-phosphate. The overexpressed enzymes were: xylose isomerase (XI), xylulokinase (XKS), ribulose-5-phosphate isomerase (R5PI), ribulose-5-phosphate epimerase (R5PE), transketolase (TKT) and transaldolase (TAL). Additionally the non-specific aldose reductase encoded by GRE3, which mediates unwanted production of xylitol [18] was deleted. Since some of the substrates of the enzymes in the PPP are not commercially available it was decided to check for overexpression via DNA arrays rather than via enzyme

activity measurements. The results listed in Table 1 further confirmed that the transcription of the target genes was successfully modified in strain RWB 212.

5 Table 1. mRNA levels of structural genes encoding xylulokinase and pentose-phosphate-pathway enzymes in the reference strain *S. cerevisiae* CEN.PK113-7D and in the engineered, xylose-isomerase-expressing strain *S. cerevisiae* RWB212. Both strains were grown in aerobic, glucose-limited chemostat cultures ($D = 0.10 \text{ h}^{-1}$). Transcript levels were determined with Affymetrix GeneChip microarrays. Data are the average \pm average deviation of the mean of analyses on three independent cultures for
10 each strain. *ACT1* (Ng and Abelson, 1980, Proc. Nat. Acad. Sci. USA. 77: 3912-3916) and *PDA1* (Wenzel et al., 1995, Nucleic. Acids Res. 23: 883-884) are included as internal standards.

Gene	Systematic name	Enzyme name	Transcript level		Fold-change
			CEN.PK113-7D	RWB212	Mutant vs WT
XylA	-	xylose isomerase	n.d.	n.d.	
XKS1	YGR194C	Xylulokinase	91 \pm 7	321 \pm 54	+ 3.5
TAL1	YLR354C	Transaldolase	574 \pm 49	959 \pm 91	+ 1.7
TKL1	YPR074C	transketolase 1	450 \pm 37	1982 \pm 79	+ 4.4
RPE1	YJL121C	D-ribulose-5-Phosphate 3-epimerase	299 \pm 24	2551 \pm 385	+ 8.5
RKI1	YOR095C	D-ribose-5-phosphate ketol-isomerase	96 \pm 8	483 \pm 64	+ 5.0
GRE3	YHR104w	aldose reductase	322 \pm 6	12 \pm 0	- 26.8
ACT1	YFL039C	Actin	408 \pm 32	366 \pm 56	NC ^a
PDA1	YER178W	E1 α subunit of pyruvate dehydrogenase complex	2901 \pm 142	3217 \pm 182	NC

n.d. = not determined (not represented on Affymetrix microarrays); ^a NC= not changed.

2.2 Physiological characterisation of the engineered strain

One of the striking properties of the engineered strain was its ability to grow anaerobically on xylose (Fig. 1) without any selective pressure being required. Anaerobic growth on xylose in mineral medium proceeded with a growth rate as high
 5 as 0.09 h^{-1} . Xylulose was not accumulated but xylitol formation, though extremely small, was detectable (Fig. 1) biomass, ethanol and glycerol yields of strain RWB 212 on xylose were comparable to those of RWB 202-AFX which was obtained via evolutionary engineering (Table 2). From Table 2 a specific xylose consumption rate of more than $1.0 \text{ g xylose/g biomass/h}$ can be calculated ($Q_s = 0,09/0,085 = 1,059 \text{ gr}$
 10 Xyl/gr X/h), compared to $345 \text{ mg xylose/g biomass/h}$ for RWB 202-AFX while a yield at least similar to the yield on glucose was obtained.

2.3 Mixed substrate utilisation

As pointed out in the introduction: economic conversion of hemicellulose hydrolysates to ethanol requires the fermentation of both glucose and xylose, preferably
 15 simultaneously. In order to verify the properties of strain RWB 212 with respect to mixed sugar utilisation, the yeast was grown in a mixture of glucose and xylose (20 g l^{-1} each). The results depicted in Fig. 2 show that both sugars were completely consumed but glucose was the preferred substrate. Xylose consumption commenced after approximately 80% of the glucose was consumed.

3. Functional expression of the *B. thetaiotaomicron* xylose isomerase in yeast

The nucleotide sequence encoding the *B. thetaiotaomicron* VPI-5482 xylose isomerase (Acc. No.'s AAO75900 or NP 809706; SEQ ID NO. 10) was cloned into a multicopy yeast expression vector to give p426GPDBtXI. This plasmid was used to transform RWB215 (*MAT α ura3-52 leu2-112 loxP-PTPI::(-266,-1)TAL1 gre3::hphMX*
 25 *pUGP_{TPI}-TKL1 pUGP_{TPI}-RPE1 KanloxP-P_{TPI}::(-?,-1)RKII*), which was further transformed with p415ADHXKS for overexpression of xylulokinase. Two independent transformants were picked and both were able to grow on minimal medium with xylose as sole carbon source and in lysates of the transformants a specific xylose isomerase activity of $140 \pm 20 \text{ U per mg protein}$ was measured, compared to about 1300 U per
 30 mg protein for the strains expressing the *Piromyces* xylose isomerase.

4. Selection of RWB 212

The strain RWB 212 (see above) was placed under selective pressure by cultivation in anaerobic chemostat cultures (duplo) with 30 g/l xylose as the carbon

source with an estimated growth rate of 0.06 h^{-1} . The selection process was monitored by determination of culture dry weight and residual xylose concentration. The initial residual xylose concentration was around 30 mM but already after 26 generations (300 hours) residual xylose concentrations had decreased to less than 15 mM (Figure 3A) and a corresponding increase in biomass concentration was also observed (Figure 3B).

From these chemostat cultures samples were taken at 530 hours and these were plated on mineral medium agar plates supplemented with 2% glucose. After 62 hours at 30°C single colonies from these plates were restreaked on fresh glucose plates. After another 72 hours incubation at 30°C, two colonies were selected (one colony originating from each separate chemostat) and used to inoculate precultures (aerobic shake flasks, mineral medium with glucose) for anaerobic fermentor batches on 20 g/l glucose and 20 g/l xylose.

From the CO₂ off-gas signals in Figure 4 it is evident that these cultures have superior xylose utilization characteristics compared to the parental strain RWB 212 and the other selection strain RWB 202-AFX. The “new” selection strain displays an increase in CO₂ production rate when consuming xylose, which is not observed in the parental strain. Figure 5 depicts the measured concentrations of carbon source and products in supernatants of these two independent batches, mainly during the xylose consumption phase. The overall (i.e. glucose + xylose phase) volumetric ethanol production rate of both experiments is higher than 0.5 g/L.hr and the first batch also demonstrates a volumetric productivity in the xylose consumption phase of higher than 0.5 g/L.hr.

5. Testing of strains RWB 204, 206, 208 and 210

The strains tested have been constructed as has been described in the patent text as well as in Kuyper et al., 2005, FEMSYR 5: 399-409. The modified genes are listed in the Table below:

Table listing of the genes overexpressed and deleted in the used strains

Strain	Overexpression	Deletion
RWB 204	TAL1	
RWB 206	TAL1	gre3
RWB 208	TAL1, TKL1	gre3
RWB 210	TAL1, TKL1, RPE1	gre3
RWB 212	TAL1, TKL1, RPE1, RKI1	gre3

After the introduction of the modifications listed in the above Table all strains were transformed with two plasmids; pAKX002 expressing the *Piromyces* xylose isomerase and p415ADHXKS a second plasmid expressing the endogenous xylulokinase. In the above article RWB 212 transformed with the two plasmids was given a separate number: RWB 217. A repeat of the transformation of RWB 212 with the two plasmids resulted in RWB 223.

After transformation single colonies were streaked on synthetic medium agar plates with glucose. From these plates shake flask cultures with synthetic medium and 20 g/l xylose were inoculated and incubated at 30°C for 48 hours. Each shake flask culture was used to inoculate an anaerobic fermentor with synthetic medium and 20 g/l xylose.

After 48 hours incubation the shake flask inoculated with RWB 204 had not grown as determined by visual inspection. All four flasks were used to inoculate one fermentor. The growth in the fermentors was monitored by measuring the CO₂ concentrations in off gas.

For reference purposes the CO₂ profiles of RWB 217 and RWB 223 (both with *TAL*, *TKL*, *RPE* and *RKI* overexpressions) are also given in figure 1. Over a period of 100 hours no significant CO₂ production could be measured in the off gas of the batches with RWB 204 and RWB 206. The growth rate determined from these CO₂ production graphs is 0.12 h⁻¹ for RWB 208, 217 and 223, for RWB 210 it was determined at 0.08 h⁻¹. From these result follows that overexpression of both transaldolase and transketolase are required for anaerobic growth on xylose. Furthermore, the additional overexpression of ribose 5-phosphate epimerase in RWB 210 decreases the growth rate on xylose. The overexpression of RPE1 probably disturbs the equilibrium between xylulose-5P, ribulose-5P and ribose-5P, hindering the activity of transketolase. Under these experimental conditions the additional overexpression of the R5P-epimerase and -isomerase does not further improve the performance of anaerobic xylose fermentation.

Table 2. Growth parameters, sugar consumption and product formation by the wild-type CEN.PK 113-7D, the selected strain RWB 202-AFX and the engineered strain RWB 212 during anaerobic batch cultivation in fermenters. Values are presented as the average and experimental deviation of two independent batch cultivations.

	CEN.PK 113-7D	RWB 202-AFX	RWB 202-AFX	RWB 212	RWB 212
Carbon source (w/v)	Glucose (2%)	Glucose (2%)	Xylose (2%)	Xylose (2%)	Glucose (2%) + Xylose (2%)
Specific growth rate (h ⁻¹)	0.34 ± 0.00	0.24 ± 0.00	0.03 ± 0.00	0.09 ± 0.00	0.25 ± 0.00 ^a
Biomass yield (g g ⁻¹)	0.099 ± 0.003	0.079 ± 0.000	0.088 ± 0.004	0.085 ± 0.002	0.074 ± 0.001
Ethanol yield ^b (g g ⁻¹)	0.40 ± 0.01	0.40 ± 0.00	0.42 ± 0.00	0.43 ± 0.00	0.43 ± 0.00
Carbon recovery ^b (%)	104.0 ± 1.1	103.7 ± 0.8	105.5 ± 0.0	105.9 ± ??	103.2 ± ??
Sugar consumed (mM)	116.1 ± 0.3	114.9 ± 0.4	137.4 ± 0.2	133.9 ± 0.1	108.5 ± 0.2 + 136.0 ± 0.3
Products:					
Biomass (g l ⁻¹)	2.07 ± 0.06	1.64 ± 0.01	1.81 ± 0.08	1.70 ± 0.04	2.97 ± 0.04
CO ₂ (mmoles l ⁻¹)	197.1 ± 3.4	196.9 ± 1.3	199.7 ± 1.5	199.9 ± 1.5	391.6 ± 0.6
Ethanol ^c (mM)	181.6 ± 3.9	180.3 ± 1.4	186.8 ± 2.2	188.5 ± 1.3	370.7 ± 0.4
Xylitol (mM)	< 0.01	< 0.01	2.76 ± 0.03	0.38 ± 0.04	0.78 ± 0.00
Xylulose (mM) ^d	< 0.01	< 0.01	7.98 ± 0.09	< 0.01	< 0.01
Glycerol (mM)	22.9 ± 0.2	24.2 ± 0.1	18.3 ± 0.3	17.8 ± 0.2	32.7 ± 0.3
Acetate (mM)	3.42 ± 0.11	6.93 ± 0.02	2.26 ± 0.16	1.40 ± 0.07	3.54 ± 0.02
Succinate (mM)	0.26 ± 0.01	0.27 ± 0.02	0.75 ± 0.00	0.39 ± 0.02	0.96 ± 0.00
Lactate (mM)	1.70 ± 0.02	1.49 ± 0.02	0.95 ± 0.02	1.46 ± 0.01	2.78 ± 0.03

5 ^a determined from the glucose consumption phase. ^b calculation based on the ethanol concentrations deduced from the CO₂ production, see Section 1.10. ^c deduced from the CO₂ production, see Section 1.10. ^d transient accumulation. This value represents the highest concentration during the mid-log phase. At the end of growth all xylulose had been reconsumed, in all other cultures the xylulose concentration remained below the
10 detection limit.

Table 3: primers used in this study

Oligo name	
P _{TAL} disA	CCTTTCCAACGAATCGTATATACTAACATGCGCGCGCTTCCTATGCATAGGCCACTAGTGGATCTG
P _{TAL} disB	AGAGAGTTGTTAGCAACCTTTTGTTCCTTTTGGAGCTGGTTCAGACATGGTGAATTCCTGTATGTG
5'TAL1	CTGACTGAGCCATATTGAGG
TAL1 intern	CACCAGTGTCTGGCAACAACG
P _{RKI} disA	TCTTGTAGAAAATTAACAACATCGTGTTACATAAACTTGGTTACGCATAGGCCACTAGTGGATCTG
P _{RKI} disB	TTGCCCAAAGATTCTAACGCATCAATTTTGGGACACCGGCAGCCATGGTGAATTCCTGTATGTG
RKI1intern	CAGCTCTCTTGGCATCCTCC
EcoRI-5'TKL1	GGAATTCATGACTCAATTCAGTACATTG
3'TKL1-XhoI	GGCCTCGAGCTTGAATGGTGTGATTCTCT
TKL1intern	CCGCCATTGGTGATCTACAG
EcoRI-5'RPEI	GGAATTCATGGTCAAACCAATTATAGC
3'RPEI-XhoI	CCGCTCGAGTTAGGCACTTACGTATCTTG
RPEIintern	GGAAGCCTTATGGAGTGTCA
P _{TP1} primer	TGGCATGTGAGATTCTCCGA
KanA	CGCACGTCAAGACTGTCAAG
KanB	TCGTATGTGAATGCTGGTCG
5'gre3::Kanlox	AAAATACTGTAATATAAATCGTAAAGGAAAATTGGAAATTTTTTCAAGCTGAAGCTTCGTACGC
3'gre3::Kanlox	TGGATTTTACTGGCTGGATCAGGCAAAAGTGGGGAATTTACCGCATAGGCCACTAGTGGATCTG
5'GRE3	CCTGGTGGAACATCCTAGAA
3'GRE3	GGATGACACCACAGGCAGAA
SpeI-5'XKS1	GACTAGTATGTTGTGTTTCAGTAATTCAG
3'XKS1-SalI	TGCAGTCGACATTTTAGATGAGAGTCTTTCC

Table 4: plasmids used in this paper

pUG6	loxP-KanMX-loxP cassette	Guldener et al. [3]
pUG6 _{TPI1}	pUG6 with the TPI1 promoter	this work
pUG6 _{TPI1-RPE1}	pUG6 with RPE1 behind the TPI1 promoter	this work
pUG6 _{TPI1-TKL1}	pUG6 with TKL1 behind the TPI1 promoter	this work
pAG32	loxP-hphMX-loxP cassette	Goldstein and McCusker [9]
PAKX002	2 μ , <i>URA3</i> , <i>Piromyces XylA</i> behind the TPI1 promoter	Kuyper et al. [20]
P415ADH	CEN, <i>LEU2</i> , <i>ADH1</i> promoter	Mumberg et al. [21]
p415ADHXKS1	CEN, <i>LEU2</i> , P _{ADH1} - <i>XKS1</i>	this work
PSH47	CEN, <i>URA3</i> , Cre recombinase behind P _{GAL1}	Guldener et al. [3]

References

1. Bruinenberg, P. M., De Bot, P. H. M, Van Dijken, J. P. and Scheffers, W. A. (1983) The role of the redox balance in the anaerobic fermentation of xylose by yeasts. Eur. J. Appl. Microbiol. Biotechnol. 18, 287-292.
5
2. Kuyper, M., Winkler, A. A., Van Dijken, J. P. and Pronk, J. T. (2004) Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle. FEMS Yeast Res. 4, 655-664.
3. Guldener, U., Heck, S., Fiedler, T., Beinhauer, J. and Hegemann, J. H. (1996) A new efficient gene disruption cassette for repeated use in budding yeast. Nucleic Acids Res. 24, 2519-2524.
10
4. Mumberg, D., Muller, R. and Funk, M. (1995) Yeast Vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene 156, 119-122.
5. Sambrook, K., Fritsch, E. F. and Maniatis, I. (1989) Molecular cloning: A laboratory manual, 2nd edn., Cold spring harbour, New York.
15
6. Verduyn, C., Postma, E., Scheffers, W. A. and Van Dijken, J. P. (1992) Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. Yeast 8, 501-517.
7. Gietz, R. D. and Woods, R. A. (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 350, 87-96.
20
8. Inoue, H., Nojima, H. and Okayama, H. (1990) High efficiency transformation of *Escherichia coli* with plasmids. Gene 96, 23-28.
9. Goldstein, A. L. and McCusker, J. H. (1999) Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. Yeast 15, 1541-1553.
25
10. Andreasen, A. A. and Stier, T. J. (1953) Anaerobic nutrition of *Saccharomyces cerevisiae*. I. Ergosterol requirement for growth in a defined medium. J. Cell Physiol. 41, 23-36.
30
11. Andreasen, A. A. and Stier, T. J. (1954) Anaerobic nutrition of *Saccharomyces cerevisiae*. II. Unsaturated fatty acid requirement for growth in a defined medium. J. Cell Physiol. 43, 271-281.

12. Van Urk, H., Mak, P. R., Scheffers, W. A. and Van Dijken, J. P. (1988) Metabolic responses of *Saccharomyces cerevisiae* CBS 8066 and *Candida utilis* CBS 621 upon transition from glucose limitation to glucose excess. *Yeast* 4, 283-291.
13. Weusthuis, R. A., Luttk, M. A., Scheffers, W. A., Van Dijken, J. P. and
5 Pronk, J. T. (1994) Is the Kluyver effect in yeasts caused by product inhibition? *Microbiology* 140, 1723-1729.
14. Verduyn, C., Postma, E., Scheffers, W. A. and Van Dijken, J. P. (1990) Physiology of *Saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures. *J. Gen. Microbiol.* 136, 395-403.
- 10 15. Piper, M. D., Daran-Lapujade, P., Bro, C., Regenber, B., Knudsen, S., Nielsen, J. and Pronk, J. T. (2002) Reproducibility of oligonucleotide microarray transcriptome analyses. An interlaboratory comparison using chemostat cultures of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 277, 37001-37008.
- 15 16. Boer, V. M., de Winde, J. H., Pronk, J. T. and Piper, M. D. (2003) The genome-wide transcriptional responses of *Saccharomyces cerevisiae* grown on glucose in aerobic chemostat cultures limited for carbon, nitrogen, phosphorus, or sulfur. *J. Biol. Chem.* 278, 3265-3274.
- 20 17. Tusher, V. G., Tibshirani, R. and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U. S. A.* 98, 5116-5121.
18. Träff, K. L., Otero Cordero, R. R., Van Zyl, W. H. and Hahn-Hägerdal, B. (2001) Deletion of the GRE3 aldose reductase gene and its influence on xylose metabolism in recombinant strains of *Saccharomyces cerevisiae* expressing the *xyIA* and *XKS1* genes. *Appl. Environ. Microbiol.* 67, 5668-5674.
- 25 19. Jeffries, T. W. and Jin, Y. S. (2004) Metabolic engineering for improved fermentation of pentoses by yeasts. *Applied Microbiology and Biotechnology* 63, 495-509.
- 30 20. Kuyper, M., Harhangi, H. R., Stave, A. K., Winkler, A. A., Jetten, M. S., De Laat, W. T., D Ridder, J. J., Op den Camp, H. J., Van Dijken, J. P. and Pronk, J. T. (2003) High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*? *FEMS Yeast Res.* 4, 69-78.

21. Mumberg, D., Muller, R. and Funk, M. (1995) Yeast Vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 156, 119-122.

DEMANDES OU BREVETS VOLUMINEUX

**LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS
COMPREND PLUS D'UN TOME.**

CECI EST LE TOME _____ DE _____

NOTE: Pour les tomes additionels, veuillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

**THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE
THAN ONE VOLUME.**

THIS IS VOLUME 1 OF 2

NOTE: For additional volumes please contact the Canadian Patent Office.

Claims

1. A eukaryotic host cell with the ability to directly isomerise xylose into xylulose, whereby the host cell comprising a genetic modification that increases the flux of the pentose phosphate pathway, and whereby the host cell has a specific xylose consumption rate of at least 346 mg xylose/g biomass/h.
5
2. A host cell according to claim 1, whereby the genetic modification comprises overexpression of at least one gene of the non-oxidative part pentose phosphate pathway.
10
3. A host cell according to claim 2, whereby the gene is selected from the group consisting of the genes encoding for ribulose-5-phosphate isomerase, ribulose-5-phosphate epimerase, transketolase and transaldolase.
15
4. A host cell according to claim 2, whereby the genetic modification comprises overexpression of at least the genes coding for a transketolase and a transaldolase.
5. A host cell according to any one of claims 2 - 4, whereby the host cell further comprises a genetic modification that increases the specific xylulose kinase activity.
20
6. A host cell according to claim 5, whereby the genetic modification comprises overexpression of a gene encoding a xylulose kinase is overexpressed.
7. A host cell according to any one of claims 2 - 6, whereby the gene that is overexpressed is endogenous to the host cell.
25
8. A host cell according to any one of claims 1 - 7, whereby the host cell comprises a genetic modification that reduces unspecific aldose reductase activity in the host cell.
30
9. A host cell according to claim 8, whereby the genetic modification reduces the expression of, or inactivates a gene encoding an unspecific aldose reductase.

10. A host cell according to claim 9, whereby the gene is inactivated by deletion of at least part of the gene or by disruption of the gene.
11. A host cell according to claims 8 or 9, whereby the expression of each gene in the
5 host cell that encodes an unspecific aldose reductase is reduced or inactivated.
12. A host cell according to any one of the preceding claims, wherein the host cell is a yeast, preferably a yeast that belongs to one of the genera: *Saccharomyces*,
Kluyveromyces, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*,
10 *Schwanniomyces*, and *Yarrowia*.
13. A host cell according to claim 12, wherein the yeast belongs to one of the species: *S. cerevisiae*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus*, and *K. fragilis*.
15
14. A host cell according to any one of claims 1 - 11, wherein the host cell is a filamentous fungus, preferably a filamentous fungus that belongs to one of the genera: *Aspergillus*, *Trichoderma*, *Hemicola*, *Acremonium*, *Fusarium*, and *Penicillium*.
- 20 15. A host cell according to any one of the preceding claims, whereby the host cell expresses one or more enzymes that confer to the host cell the ability to produce at least one fermentation product selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, β -lactam antibiotics and cephalosporins.
- 25
16. A process for producing ethanol, whereby the process comprises the steps of:
(a) fermenting a medium containing a source of xylose with a modified host cell as defined in any one of claims 1 - 14, whereby the host cell ferments xylose to ethanol, and optionally,
30 (b) recovery of the ethanol.
17. A process according to claim 16, whereby the medium also contains a source of glucose.

18. A process according to claims 16 or 17, whereby the volumetric ethanol productivity is at least 0.5 g ethanol per litre per hour.
- 5 19. A process according to claims any one of claims 16 - 18, whereby the ethanol yield is at least 50 %.
- 10 20. A process for producing a fermentation product selected from the group consisting of lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, β -lactam antibiotics and cephalosporins, whereby the process comprises the steps of:
- (a) fermenting a medium containing a source of xylose with a modified host cell as defined in claim 15, whereby the host cell ferments xylose to the fermentation product, and optionally,
- 15 (b) recovery of the fermentation product.
21. A process according to claim 20, whereby the medium also contains a source of glucose.

20

25

30

Fig 1a

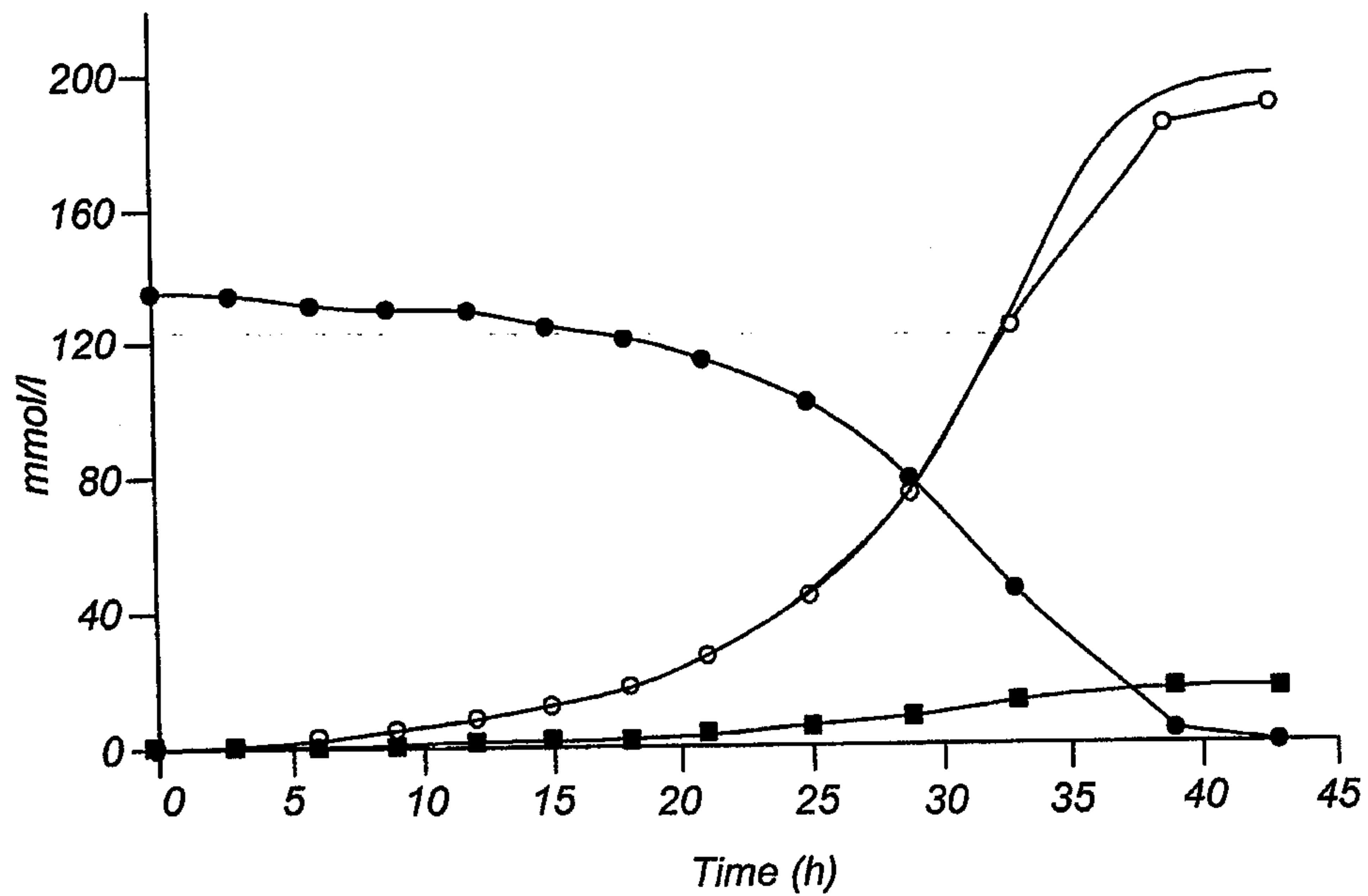


Fig 1b

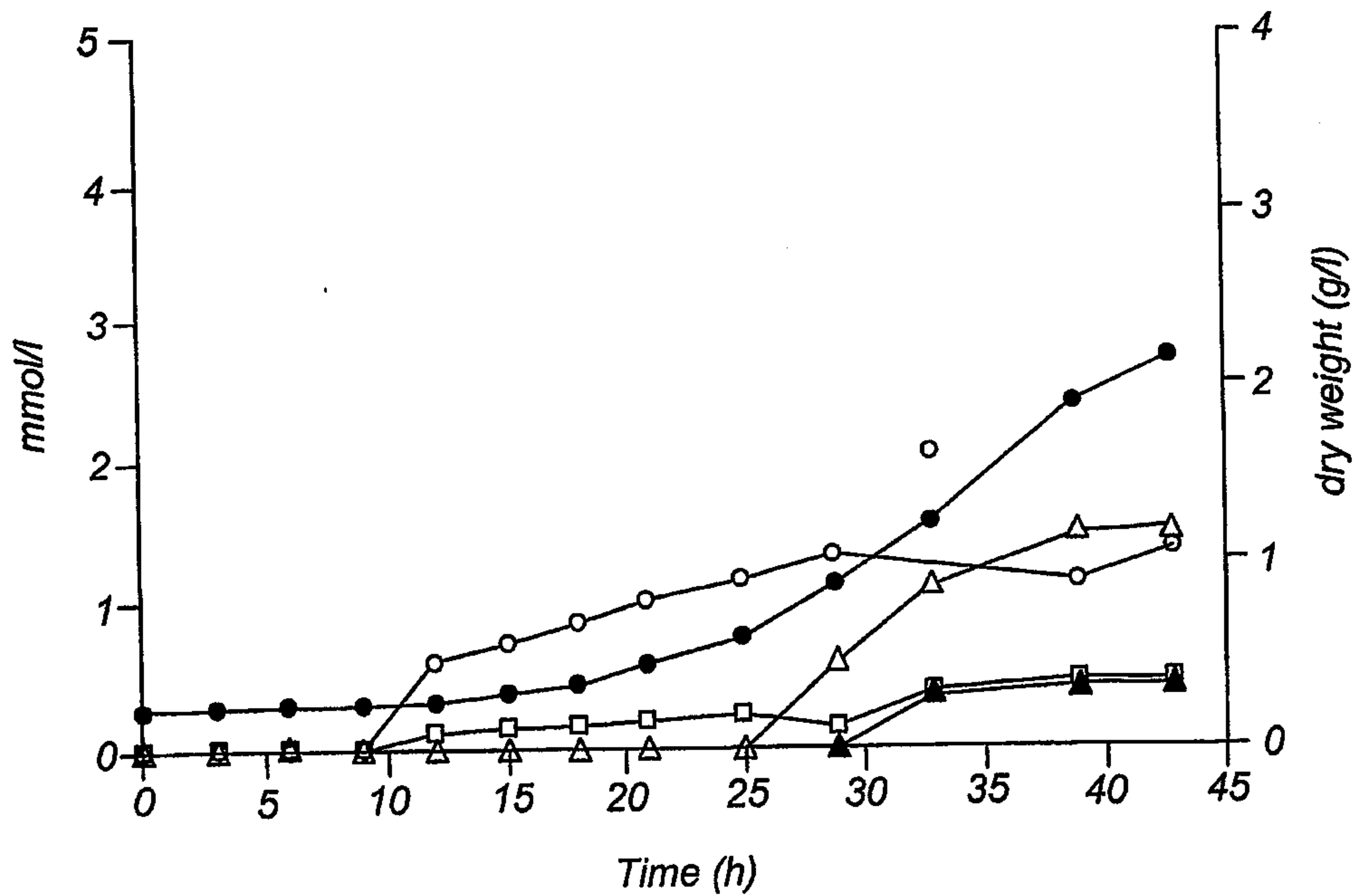


Fig 2a

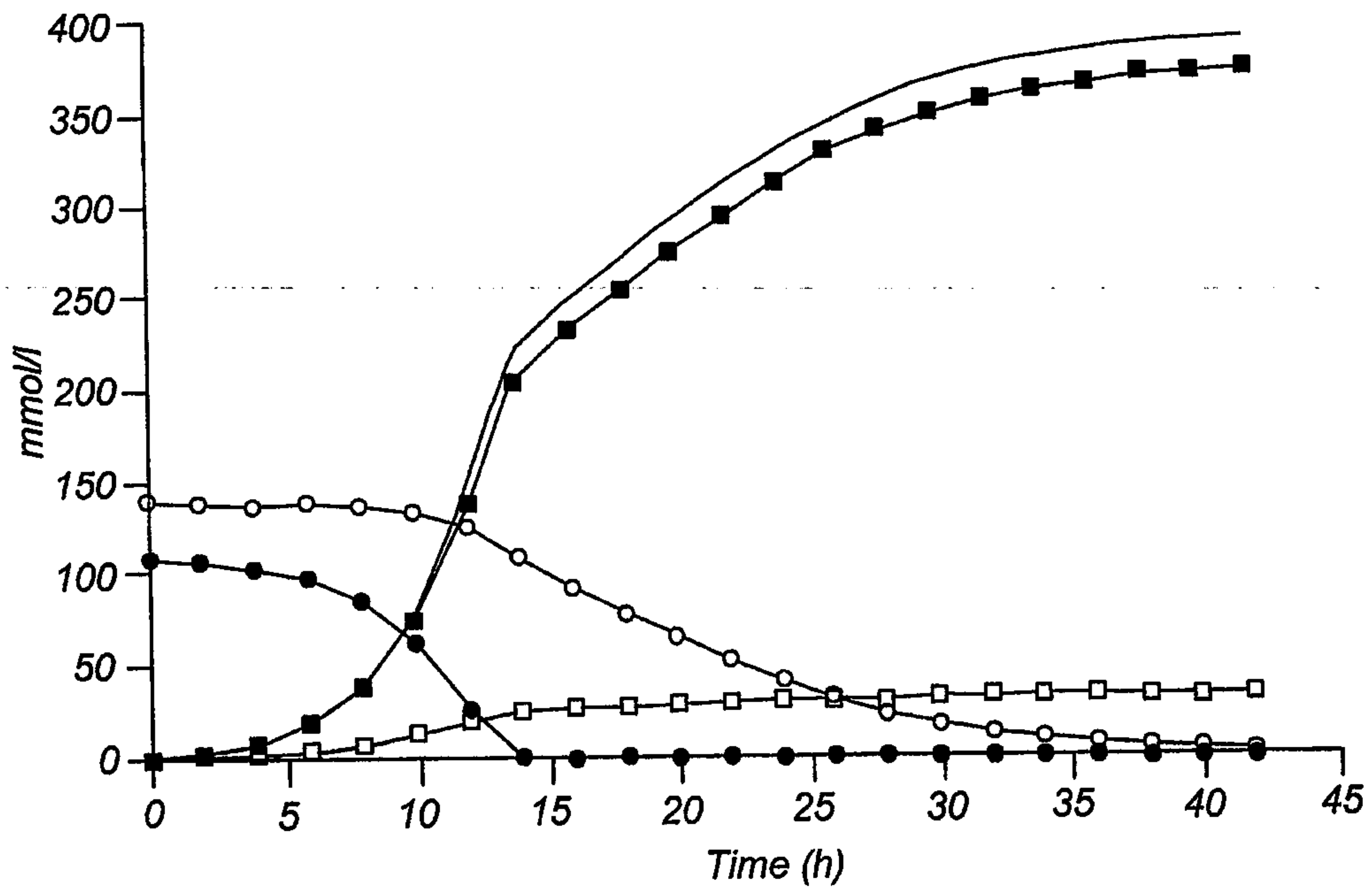


Fig 2b

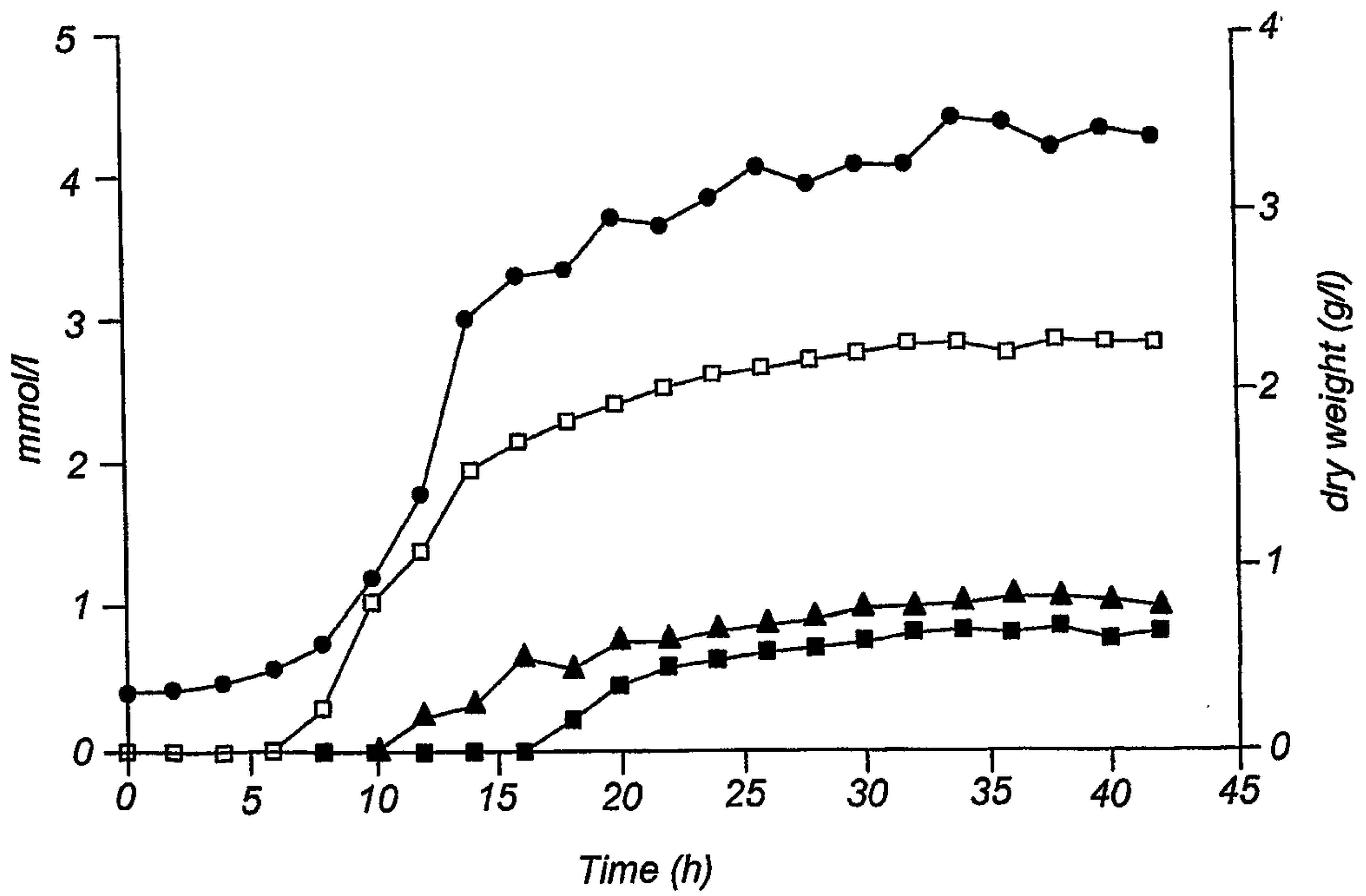


Fig 3a

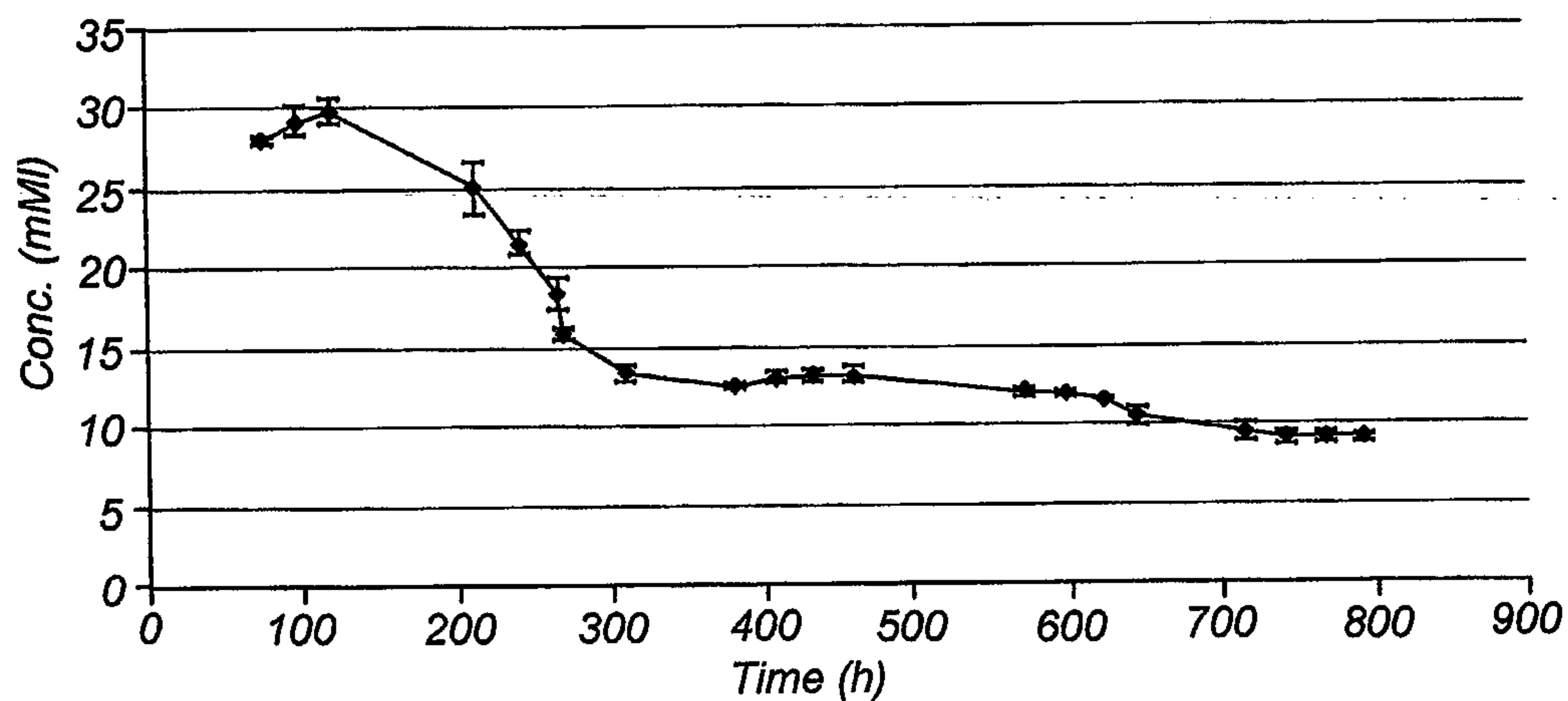


Fig 3b

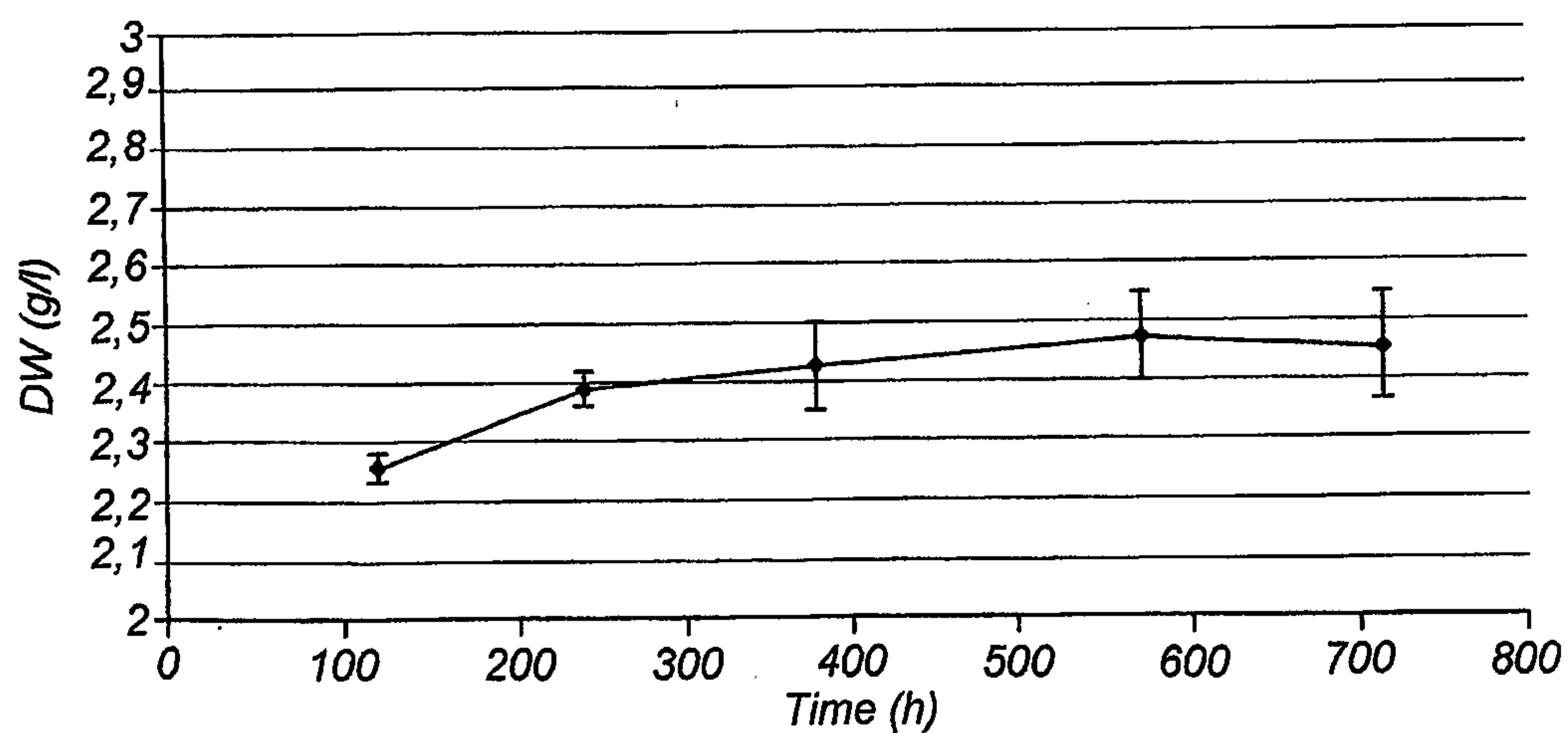


Fig 4

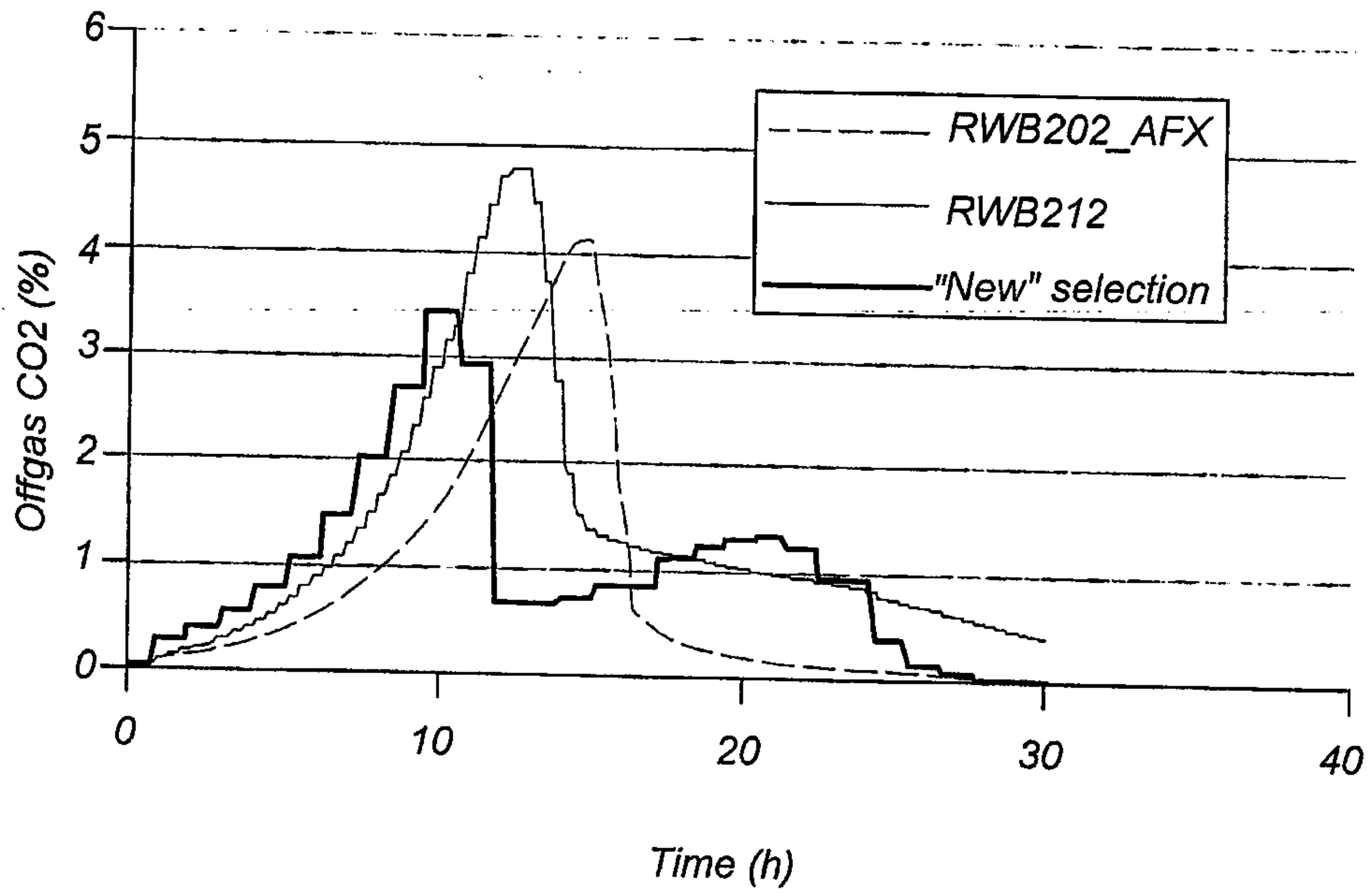


Fig 5a

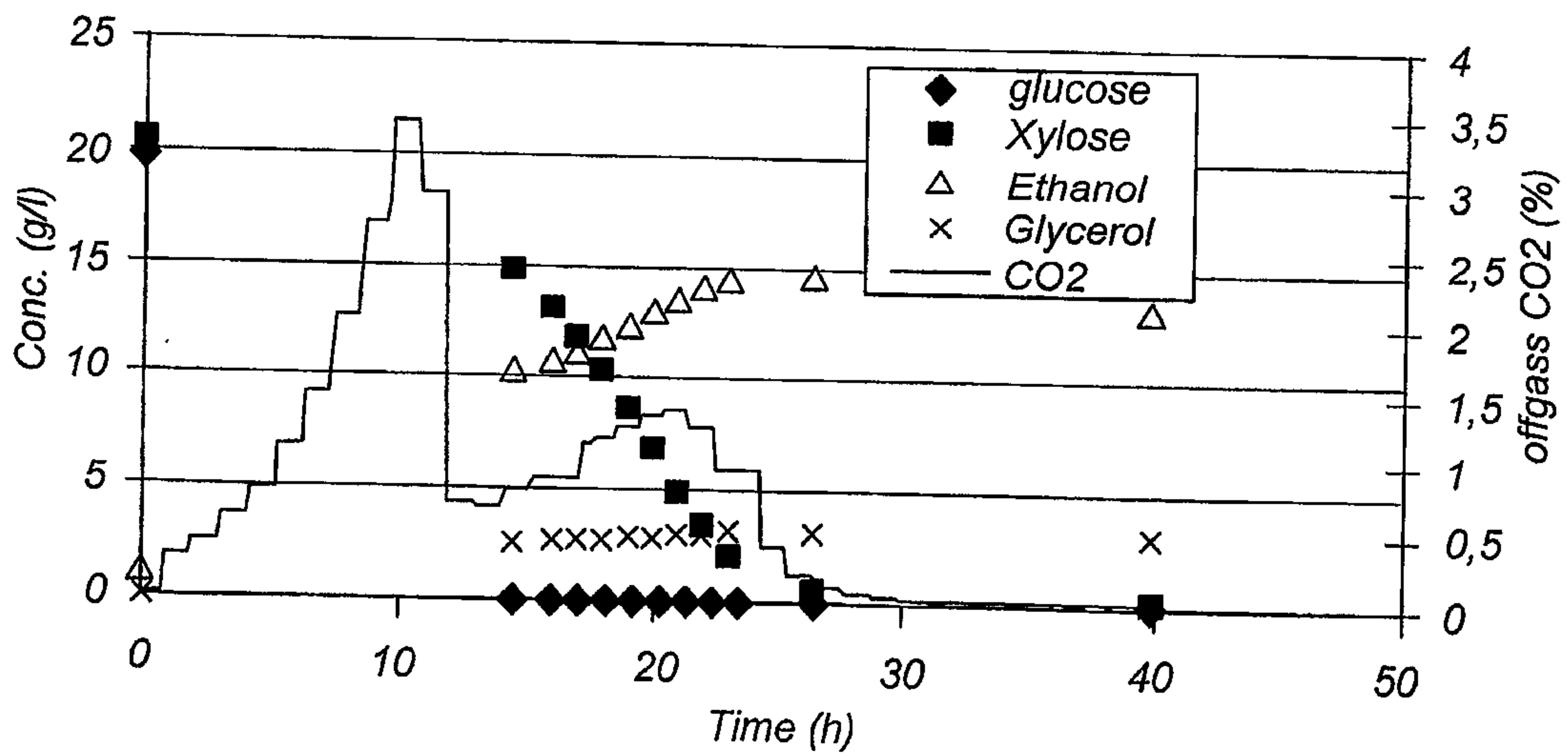


Fig 5b

