

Cloning, Expression and Characterization of UDP-glucose Dehydrogenases

Márcia R Couto, Joana L Rodrigues* and Lígia R Rodrigues

Centre of Biological Engineering, University of Minho, Braga 4710-057, Portugal

*Corresponding author: joanarodrigues@ceb.uminho.pt

Supplementary Information

Table S1. Primers used in this study for gene amplification, colony PCR and sequencing. Histidine Tag sequence is underlined.

Primers	Template	Sequence (5' → 3')	Restriction enzyme
<i>Gene amplification for cloning</i>			
Zm_Fw	<i>Zymomonas mobilis</i> gDNA	AAAGGATCCAAAAAATG <u>CATCACCATCACCATC</u> ACCGCATTACCA	<i>Bam</i> HI
Zm_Rv		TGATCGGTTCC	
Lbj_Fw	<i>Lactobacillus johnsonii</i> gDNA	AAAGGATCCAAAAAATG <u>CATCACCATCACCATC</u> ACAATTCATATA	<i>Bam</i> HI
Lbj_Rv		AAATTGCAGTTG	
Zm_prsf_Fw	p426GPD_ZmUGD	AAAGGATCCAAAAAATG <u>CATCACCATCACCATC</u> ACTTTGAGATTA	<i>Eco</i> RI
Zm_prsf_Rv		AAAAAATCTGC	
Lbj_prsf_Fw	pGEM_LbjUGD	AAAGGATCCAAAAAATG <u>CATCACCATCACCATC</u> ACTTTGAGATTA	<i>Eco</i> RI
Lbj_prsf_Rv		AAAAAATCTGC	
Ch_Fw	pUC19_ChUGD	AAAGGATCCAAAAAATG <u>CATCACCATCACCATC</u> ACTTTGAGATTA	<i>Bam</i> HI
Ch_Rv		AAAAAATCTGC	
<i>Colony PCR and Sequencing</i>			
M13_Fw	pGEM®-T Easy and derivatives	GTAAAACGACGGCCAGT	
M13_Rv		GCGGATAACAATTTCACACAGG	
pRSFDuet_Fw	pRSFDuet-1 and derivatives	GGATCTCGACGCTCTCCCTT	
T7 terminator_Rv		CTAGTTATTGCTCAGCGGT	
PGK1p_Fw	pSP-GM1 and derivatives	GTGATCTCCAGAGCAAAGT	
CYCt_Rv		CTTCGAGCGTCCCAAAC	
p426seq_Fw	p426GPD and derivatives	ACTAAAGGGAACAAAAGCTG	
CYC 1_Rv		ACTTCAGGTTGTCTAACTCC	

Table S2. Gene sequence of UGD from *Capra hircus* codon optimized (NZYTech) for *Saccharomyces cerevisiae*

>ugd1 (UDP-glucose 6-dehydrogenase) from <i>C. hircus</i> (1485 bp; GC%: 41.89; CAI: 0.74)
ATGTTTGAGATTAAAAAATCTGCTGCATCGGTGCGGGTTATGTCGGTGGTCCTACTTGT AGCGTTATTGCACACATGTGTCCTGAAATTCGTGTCACCGTCGTTGACATCAATGAGTCA AGAATAAATGCTTGGAATAGTCCGACTCTACCAATTTATGAACCCGGTTTGAAAGAAGT AGTTGAATCTTGACAGAGGTAAGAAGTTGTTTTTCTCTACAAACATCGACGACGCGATAAA GGAAGCTGATCTGGTATTCATCTCTGTTAATACACCCACCAAAACGTACGGTATGGGCAA GGGTAGAGCTGCTGATCTAAAGTACATCGAAGCGTGTGCACGTCTGAATAGTTCAGAACT CTCATGGTTACAAGATCGTGACAGAAAAGTCCACAGTTCCAGTTAGAGCAGCTGAATCC ATAAGGAGGATCTTCGATGCTAACACAAAGCCAAACTTGAAGTTACAGGTATTGTCAAA CCCGGAATTTTGTAGCTGAAGGGACGGCCATCAAGGATTTGAAGAACCCAGACCGAGTTC TAATCGGTGGTGACGAGACACCTGAAGGTCAAAGAGCTGTCCAAGCTCTCTGTGCAGTTT ACGAGCATTGGGTGCCCCGTGAAAAAATTCTTACAACCAACACTTGGTCTTCAGAGTTGT CGAAACTAACCGCTAATGCATTTTGTAGCTCAAAGAATTTCTAGCATCAATTCAATTAGTG CCTTGTGTGAAGCAACCGGCGCAGATGTAGAGGAAGTCGCTACAGCTATTGGAATGGAT CAAAGAATAGGTTCTAAATTTCTAAAAGCTAGCGTAGGATTTGGTGGTTCTTGCTTTCAA AAGGATGTCCTTAATTTAGTGTATTTGTGTGAAGCCCTTAATTTACCAGAAGTCGCTAGAT ACTGGCAACAGGTCATAGATATGAACGACTATCAACGTAGAAGATTTCGCTTCGAGGATA ATTGATTCTCTATTTAACACAGTTACTGATAAAAAGATTGCCATATTGGGATTTCGCCTTTA AGAAAGACACTGGCGACACAAGAGAGTCATCATCAATATACATTTCTAAGTATTTGATG GATGAAGGTGCTCACTTGCATATTTATGATCCAAAGGTGCCACGTGAACAGATTGTGGTA GATTTATCGCATCCCGGCGTTAGCAAGGATGATCAAGTTGCTCGTCTTGTGACAATCTCC AAAGACCCATACGAAGCATGTGATGGTGCTCACGCAGTGGTTATTTGTACAGAATGGGA CATGTTTAAAGAACTGGATTACGAGAGAATACACAAAAAATGTTGAAGCCCGCTTTTA TTTTTGATGGTAGGAGAGTGTTGGACGGGCTACATAACGAATTGCAAACCATAGGCTTTC AAATTGAGACAATTGGGAAGAAGGTTTCCAGTAAAAGAATCCCTTATGCGCCGTCAGGG GAGATACCAAAGTTTAGTCTACAAGACATGCCGAATAAGAAACCGAGAGTATAG

Table S3. Results of blastp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the protein domains

Query	Organism	Protein	Access Number	Identity (%)
ZmUGD	<i>Sinorhizobium meliloti</i> 1021	UGD	O54068.2	67.05
	<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	UGD (YwqF)	P96718.1	48.51
	<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	UGD (TuaD)	O32271.1	46.67
	<i>Pseudomonas aeruginosa</i> PAO1	UGD	O86422.2	45.86
	<i>Rickettsia bellii</i> RML369-C	UGD	Q1RKf8.1	44.93
	<i>Methanocaldococcus jannaschii</i> DSM 2661	UGD	Q58454.1	44.04
	<i>Rickettsia conorii</i> str. Malish 7	UGD	Q92GB1.1	44.01
	<i>Rickettsia felis</i> URRWXCα2	UGD	Q4UK39.1	43.91
	<i>Rickettsia typhi</i> str. Wilmington	UGD	Q68VX0.1	43.78
	<i>Rickettsia prowazekii</i> str. Madrid E	UGD	O05973.1	43.09
	<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	Putative UGD (YtcA)	O34862.1	40.50
	<i>Haloferax volcanii</i> DS2	UGD (AglM)	D4GYH5.1	39.00
	<i>Drosophila melanogaster</i>	UGD	O02373.1	37.53
	<i>Pseudomonas aeruginosa</i> PAO1	GMD	P11759.2	36.91
	<i>Homo sapiens</i>	UGD	O60701.1	36.85
	<i>Pongo abelii</i>	UGD	Q5R7B3.1	36.85
	<i>Bos taurus</i>	UGD	P12378.2	36.64
	<i>Mus musculus</i>	UGD	O70475.1	36.64
	<i>Oryza sativa</i> Japonica Group	UGD1	Q75GS4.1	36.46
	<i>Glycine max</i>	UGD1	Q96558.1	36.23
	<i>Gallus gallus</i>	UGD	Q5F3T9.1	36.23
	<i>Oryza sativa</i> Japonica Group	UGD2	B7F958.1	36.09
	<i>Rattus norvegicus</i>	UGD	O70199.1	36.01
	<i>Arabidopsis thaliana</i>	UGD2	Q9LIA8.1	35.96

<i>Arabidopsis thaliana</i>	UGD3	Q9LF33.1	35.73
<i>Arabidopsis thaliana</i>	UGD1	Q9FZE1.1	35.39
<i>Caenorhabditis elegans</i>	UGD	Q19905.1	35.36
<i>Arabidopsis thaliana</i>	UGD4	Q9FM01.1	35.31
<i>Azotobacter vinelandii</i>	GMD	P51585.1	35.08

Table S3. Results of blastp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the protein domains (*continuation*)

Query	Organism	Protein	Access Number	Identity (%)
ZmUGD (cont.)	<i>Oryza sativa</i> Japonica Group	UGD5	Q2QS13.1	35.02
	<i>Pseudomonas syringae</i> pv. <i>tomato</i> str. DC3000	GMD	Q887P8.1	34.99
	<i>Oryza sativa</i> Japonica Group	UGD3	Q9AUV6.1	34.75
	<i>Oryza sativa</i> Japonica Group	UGD4	Q2QS14.1	34.53
	<i>Pseudomonas putida</i> KT2440	GMD	Q88NC4.1	33.18
	<i>Pseudomonas syringae</i> pv. <i>syringae</i>	GMD	P59793.1	32.60
	<i>Escherichia coli</i> K5	UGD	Q47329.1	32.39
	<i>Pseudomonas savastanoi</i> pv. <i>phaseolicola</i>	GMD	O07299.1	31.56
	<i>Escherichia coli</i> K-12	UAMAD	P27829.4	31.01
	<i>Escherichia coli</i> O8:K40	UGD	Q33952.1	30.79
	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Typhi</i>	UAMAD	Q8Z389.1	30.66
	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar	UAMAD	Q9L6R4.1	30.39
	<i>Typhimurium</i> str. LT2			
	<i>Streptococcus pneumoniae</i>	UGD	Q57346.2	29.92
	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar	UGD	Q04873.1	29.86
	<i>Typhimurium</i> str. LT2			
	<i>Shigella flexneri</i>	Putative UGD	P37791.1	29.56
	<i>Escherichia coli</i> O157:H7	UGD	Q7DBF9.1	29.21
	<i>Escherichia coli</i> CFT073	UGD	Q8FG45.1	29.21
	<i>Escherichia coli</i> K-12	UGD	P76373.1	29.21

<i>Escherichia coli</i> O111:H-	UGD	Q04872.1	29.13
<i>Methanococcus maripaludis</i> S2	UAMAD	Q6LZC3.1	28.70
<i>Streptococcus pyogenes</i>	UGD	P0C0F4.1	28.57
<i>Streptococcus pyogenes</i> serotype M1	UGD	P0C0F5.1	28.57
<i>Methanococcus maripaludis</i> C7	UAMAD	A6VK13.1	28.47
<i>Streptococcus pyogenes</i> SSI-1	UGD	P0DG69.1	28.30
<i>Streptococcus pyogenes</i> MGAS315	UGD	P0DG68.1	28.30
<i>Methanococcus maripaludis</i> C5	UAMAD	A4FY94.1	28.02
<i>Ralstonia solanacearum</i> GMI1000	NAGAAD	P58591.1	27.87

Table S3. Results of blastp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the protein domains (*continuation*)

Query	Organism	Protein	Access Number	Identity (%)
ZmUGD (cont.)	<i>Yersinia pestis</i>	UAMAD	Q8ZAE4.1	27.85
	<i>Methanocaldococcus jannaschii</i> DSM 2661	UAMAD	Q57871.1	27.52
	<i>Ralstonia solanacearum</i>	NAGAAD	Q45410.1	27.17
	<i>Methanococcus vannielii</i> SB	UAMAD	A6USK4.1	26.90
	<i>Methanococcus aeolicus</i> Nankai-3	UAMAD	A6UU98.1	26.09
	<i>Pseudomonas aeruginosa</i> PAO1	UAGAD	G3XD94.1	25.74
LbjUGD	<i>Escherichia coli</i> O111:H-	UGD	Q04872.1	61.80
	<i>Escherichia coli</i> O8:K40	UGD	Q33952.1	61.56
	<i>Escherichia coli</i> CFT073	UGD	Q8FG45.1	60.34
	<i>Escherichia coli</i> O157:H7	UGD	Q7DBF9.1	60.10
	<i>Escherichia coli</i> K-12	UGD	P76373.1	59.85
	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar	UGD	Q04873.1	59.61
	<i>Typhimurium</i> str. LT2			
	<i>Streptococcus pyogenes</i> serotype M1	UGD	P0C0F5.1	57.77
	<i>Streptococcus pyogenes</i> WF50	UGD	P0C0F4.1	57.77
	<i>Streptococcus pneumoniae</i>	UGD	Q57346.2	57.66
	<i>Escherichia coli</i> K5	UGD	Q47329.1	57.63

<i>Streptococcus pyogenes</i> MGAS315	UGD	P0DG68.1	57.52
<i>Streptococcus pyogenes</i> SSI-1	UGD	P0DG69.1	57.52
<i>Shigella flexneri</i>	Putative	P37791.1	57.42
	UGD		
<i>Methanocaldococcus jannaschii</i> DSM 2661	UGD	Q58454.1	32.85
<i>Rickettsia prowazekii</i> str. Madrid E	UGD	Q05973.1	30.42
<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	UGD (YwqF)	P96718.1	30.03
<i>Rickettsia bellii</i> RML369-C	UGD	Q1RKf8.1	30.02
<i>Rickettsia typhi</i> str. Wilmington	UGD	Q68VX0.1	29.71
<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	UGD (TuaD)	Q32271.1	29.64
<i>Rickettsia felis</i> URRWXCal2	UGD	Q4UK39.1	29.10
<i>Sinorhizobium meliloti</i> 1021	UGD	Q54068.2	29.06
<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	Putative	Q34862.1	28.76
	UGD (YtcA)		
<i>Rickettsia conorii</i> str. Malish 7	UGD	Q92GB1.1	28.46

Table S3. Results of blastp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the protein domains (*continuation*)

Query	Organism	Protein	Access Number	Identity (%)
<i>LbjUGD</i>	<i>Azotobacter vinelandii</i>	GMD	P51585.1	28.39
(cont.)	<i>Pseudomonas syringae</i> pv. <i>Syringae</i>	GMD	P59793.1	28.00
	<i>Pseudomonas syringae</i> pv. <i>tomato</i> str. DC3000	GMD	Q887P8.1	27.93
	<i>Pseudomonas savastanoi</i> pv. <i>phaseolicola</i>	GMD	Q07299.1	27.76
	<i>Arabidopsis thaliana</i>	UGD4	Q9FM01.1	27.11
	<i>Caenorhabditis elegans</i>	UGD	Q19905.1	27.03
	<i>Pseudomonas aeruginosa</i> PAO1	GMD	P11759.2	26.77
	<i>Bacillus licheniformis</i> DSM 13 = ATCC 1458	GPD	Q65I16.1	26.53
	<i>Rattus norvegicus</i>	UGD	Q70199.1	26.51
	<i>Homo sapiens</i>	UGD	Q60701.1	26.51
	<i>Bos taurus</i>	UGD	P12378.2	26.51

	<i>Pseudomonas aeruginosa</i> PAO1	UGD	O86422.2	26.37
	<i>Haloferax volcanii</i> DS2	UGD (AglM)	D4GYH5.1	26.32
	<i>Pongo abelii</i>	UGD	Q5R7B3.1	26.25
	<i>Mus musculus</i>	UGD	O70475.1	26.25
	<i>Pseudomonas putida</i> KT2440	GMD	Q88NC4.1	25.85
	<i>Drosophila melanogaster</i>	UGD	O02373.1	25.39
	<i>Staphylococcus aureus</i>	UGD/GMD superfamily (CapL)	P39861.1	25.24
ChUGD	<i>Bos taurus</i>	UGD	P12378.2	99.80
	<i>Homo sapiens</i>	UGD	O60701.1	98.38
	<i>Pongo abelii</i>	UGD	Q5R7B3.1	98.18
	<i>Mus musculus</i>	UGD	O70475.1	96.36
	<i>Rattus norvegicus</i>	UGD	O70199.1	96.15
	<i>Gallus gallus</i>	UGD	Q5F3T9.1	92.51
	<i>Drosophila melanogaster</i>	UGD	O02373.1	68.55
	<i>Caenorhabditis elegans</i>	UGD	Q19905.1	65.87
	<i>Glycine max</i>	UGD1	Q96558.1	61.75
	<i>Arabidopsis thaliana</i>	UGD4	Q9FM01.1	61.75

Table S3. Results of blastp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the protein domains (*continuation*)

Query	Organism	Protein	Access Number	Identity (%)
ChUGD	<i>Oryza sativa</i> Japonica Group	UGD5	Q2QS13.1	61.44
(cont.)	<i>Arabidopsis thaliana</i>	UGD2	Q9LIA8.1	61.32
	<i>Arabidopsis thaliana</i>	UGD1	Q9FZE1.1	61.31
	<i>Arabidopsis thaliana</i>	UGD3	Q9LF33.1	61.11
	<i>Oryza sativa</i> Japonica Group	UGD4	Q2QS14.1	61.02
	<i>Oryza sativa</i> Japonica Group	UGD2	B7F958.1	60.98
	<i>Oryza sativa</i> Japonica Group	UGD3	Q9AUV6.1	60.81

<i>Oryza sativa</i> Japonica Group	UGD1	Q75GS4.1	59.75
<i>Sinorhizobium meliloti</i> 1021	UGD	O54068.2	37.96
<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	UGD (YwqF)	P96718.1	35.42
<i>Methanocaldococcus jannaschii</i> DSM 2661	UGD	Q58454.1	35.38
<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	UGD (TuaD)	O32271.1	34.77
<i>Haloferax volcanii</i> DS2	UGD (AglM)	D4GYH5.1	34.68
<i>Pseudomonas aeruginosa</i> PAO1	UGD	O86422.2	32.70
<i>Rickettsia conorii</i> str. Malish 7	UGD	Q92GB1.1	32.54
<i>Rickettsia felis</i> URRWXCal2	UGD	Q4UK39.1	32.47
<i>Rickettsia typhi</i> str. Wilmington	UGD	Q68VX0.1	32.47
<i>Rickettsia prowazekii</i> str. Madrid E	UGD	O05973.1	32.40
<i>Rickettsia bellii</i> RML369-C	UGD	Q1RKF8.1	32.26
<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	Putative	O34862.1	31.29
	UGD (YtcA)		
<i>Pseudomonas syringae</i> pv. <i>syringae</i>	GMD	P59793.1	29.17
<i>Pseudomonas syringae</i> pv. <i>tomato</i> str. DC3000	GMD	Q887P8.1	29.09
<i>Pseudomonas savastanoi</i> pv. <i>phaseolicola</i>	GMD	O07299.1	28.91
<i>Azotobacter vinelandii</i>	GMD	P51585.1	28.68
<i>Pseudomonas aeruginosa</i> PAO1	GMD	P11759.2	27.98
<i>Shigella flexneri</i>	Putative	P37791.1	27.82
	UGD		
<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar	UGD	Q04873.1	27.72
<i>Typhimurium</i> str. LT2			

Table S3. Results of blastp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the protein domains (*continuation*)

Query	Organism	Protein	Access Number	Identity (%)
ChUGD	<i>Methanococcus maripaludis</i> C7	UAMAD	A6VK13.1	27.47
(cont.)	<i>Escherichia coli</i> O8:K40	UGD	O33952.1	27.37

<i>Methanococcus maripaludis</i> S2	UAMAD	Q6LZC3.1	27.25
<i>Escherichia coli</i> O111:H-	UGD	Q04872.1	27.25
<i>Methanocaldococcus jannaschii</i> DSM 2661	UAMAD	Q57871.1	27.21
<i>Pseudomonas putida</i> KT2440	GMD	Q88NC4.1	27.08
<i>Methanococcus aeolicus</i> Nankai-3	UAMAD	A6UU98.1	26.87
<i>Methanococcus vannielii</i> SB	UAMAD	A6USK4.1	26.82
<i>Escherichia coli</i> K5	UGD	Q47329.1	26.53
<i>Escherichia coli</i> O157:H7	UGD	Q7DBF9.1	26.36
<i>Methanococcus maripaludis</i> C5	UAMAD	A4FY94.1	26.18
<i>Escherichia coli</i> CFT073	UGD	Q8FG45.1	26.09
<i>Escherichia coli</i> K-12	UGD	P76373.1	26.09
<i>Pyrococcus horikoshii</i> OT3	UAMAD	Q59284.1	25.74
<i>Streptococcus pneumoniae</i>	UGD	Q57346.2	25.68
<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Typhi</i>	UGD/GMD superfamily (VipA/TviB)	Q04972.1	23.79

Enzyme abbreviations: UGD, uridine diphosphate(UDP)-glucose 6-dehydrogenase; GMD, guanosine diphosphate(GDP)-mannose 6-dehydrogenase; UAMAD, UDP-N-acetyl-D-mannosamine dehydrogenase; UAGAD, UDP-N-acetylglucosamine 6-dehydrogenase; and NAGAAD, nucleoside diphosphate(NDP)-N-acetyl-D-galactosaminuronic acid dehydrogenase.

NAD-binding domain		
Lajo	-MNSYKIAVAGTGYVG---LSLATLLSQHNEVTAVDIVPEKVDLINNHKSPFIVDKIEIDF	56
Cahi	MFEIKKICCIAGAGYVGGPTCSVIAHMCPEIRVTVVDINESRINAWNPTLPITYEPGLKEV	60
Zymo	----MRITMTIGSGYVG---LVSGACFSDFGHNVICVDHDQEKIALLQKGVMPITYEPGLADM	54
	:* :*:*** *	
NAD-binding domain		
Lajo	LANKD--LNLKATDAESAYKDAADFVIIATPTNYDSE---KNFFDTSAVEAVIKLVKQY	110
Cahi	VESCR-GKNLFFSTNIDDAIKEADLVFISVNTPTKTYGMKGRAADLKYEACARRIVQN	119
Zymo	VANNVKAGRLSFSNDLASSVKNADAVFIAVGTPSRRG---DGHADLSYVFAAAKEIAAS	110
	: . :.* :.: :.*:*** *:*:.. *	
NAD-binding domain		
Lajo	NPQA-IMIISTIPVGYTNHVRKEFE-----TENIIFSPEFLREGHALYDNLHPSRIV	162
Cahi	SHGYKIVTEKSTVPVRAAESIRRIFDANTKPNLNLQVLSNPEFLAEGTAIKDLKNPDRVL	179
Zymo	VSDNTVVVDKSTVPVGTGDEVERIIREIAPDK-KIIVASNPEFLREGAALIGDFKRPDRIV	169
	: : ***:*** : : : : . : .**** ** *: * :.*:..	
NAD-binding domain		Central domain
Lajo	VGTDK---ENTKLVEAAHTFASLLQEGAEDKDIATLFGMGTAEAVKLFANTYLALRVSY	219
Cahi	IGGETPEQGQRAVQALCAVYEHWP-----REKILTNTWSSELSKLTANAFLAQRIS	233
Zymo	IGTENP---EAREIMQEIYRPLYLN-----QSPILFTSRRSALIKYASNAFLATKITF	220
	:* : :	: . :.* ** *:***: : :
Central domain		
Lajo	FNELDTYAESKGLNTQEIIDGVGLDPRIGSHYNNPSFGYGGYCLPKDTKQLLANYKDV--	277
Cahi	INSISALCEATGADVEEVATAIGMDQRIGSKFLKASVFGGSCFQKDVNLVYLCEALNL	293
Zymo	INEMADLCEAVGGDVQDIARGIGADNRIGSKFLHAGPGYGGSCFPKDTLALLKTAENF--	278
	:*. : .*: * :.: : .*: * ****: : .*.*** *: **.* : .	
Central domain		
Lajo	--PENLIEAIVKSNDRKDFIADQVLSKAGYYDYDDNNTYDPSEEQVITGVYRLTMKSN	335
Cahi	PEVARYWQQVIDMNDYQRRRFASRIIDSLFNT-----VTDKKIAILGFQAFKKD	341
Zymo	ATPLRVVESVVTANNIRKRAMARKVLQAIQKD-----MRGQKIGLLGLTFKPN	326
	. : : : * : : : * : : .	:.*: :*** :
UDP-binding domain		
Lajo	SDNFRQSSVQGVMKRVKAKGAKIIFEPFLENGSTF-----	371
Cahi	TGDTRESSSIYISKYLMDEGAHLHIYDPKVPREQIVVDLSHPGVSKDDQVARLVITSKDP	401
Zymo	TDDMRDAPSLALIQGLEDAIIHAYDEGMTQ---ARPLLPNVIF-----EDEF	373
	:. : *: : : : : ** : : *	
UDP-binding domain		
Lajo	-----FGSEVVNDLDEFKNKSDAIIAN-----RYNTILDDVKDKVYTRDIFRKD	415
Cahi	YEACDGAHAVVICTEWDMFKELDYERIHKMKLPAFIFDGRRLVDGLHNELQTIGFQIET	461
Zymo	YKVAEKAELVIVTEWDAFRALDLKRIAGLLKTP-LLLDLRNIYTE--EEAKKAGLVYHG	430
	. : : * *: . *	: : : : : . : . : .
Lajo	-----	415
Cahi	IGKKVSSKRIPYAPSGEIPKFSLQDMPNKKPRV	494
Zymo	IGRPNAG-----	437

Figure S1. Multiple amino acid sequence alignment of the three uridine diphosphate (UDP)-glucose dehydrogenases used in this study generated by Clustal Omega (1.2.4) (EMBL-EBI). Species are abbreviated as the first two letters of the genus followed by the two first letters of the species: Lajo, *Lactobacillus johnsonii*; Cahi, *Capra hircus*; and Zymo, *Zymomonas mobilis*. An asterisk (*) marks the residues that are conserved in the three sequences; a colon (:) indicates conservation between groups of strongly similar properties (scoring > 0.5 in the Gonnet Point Accepted Mutation (PAM) 250

matrix); and a period (.) indicates conservation between groups of weakly similar properties (scoring ≤ 0.5 in the Gonnet PAM 250 matrix). Conserved amino acids within 63 members of UDP-glucose 6-dehydrogenase/ guanosine diphosphate (GDP)-mannose 6-dehydrogenase (UGD/GMD) superfamily are highlighted in grey while conserved sites within only UGD enzymes are highlighted in yellow. The nicotinamide adenine dinucleotide (NAD)-binding domain contains a characteristic dinucleotide binding Rossman fold, indicated in a blue box, composed by a conserved glycine-rich phosphate-binding loop with the pattern GXGXXG linking the first β -sheet to the following α -helix that, in the case of UGD/GMD dehydrogenase enzymatic family, has a strictly conserved tyrosine so that the region becomes GXGYXG. The central domain is a long α -helical region that serves as the dimerization interface for these enzymes. The C-terminal substrate-binding domain forms an incomplete dinucleotide binding fold, primarily responsible for binding the UDP moiety of the UDP-sugar in a deep pocket. The described active site of UGD contains residues from the three domains that are highly conserved within the superfamily and include: N-terminal threonine (T₁₂₂), central α -helix lysine, (K₂₀₇), and asparagine, (N₂₁₁) and C-terminal contributes with cysteine (C₂₆₃) and aspartic acid (D₂₆₇) (reference sites from *Z. mobilis* sequence)

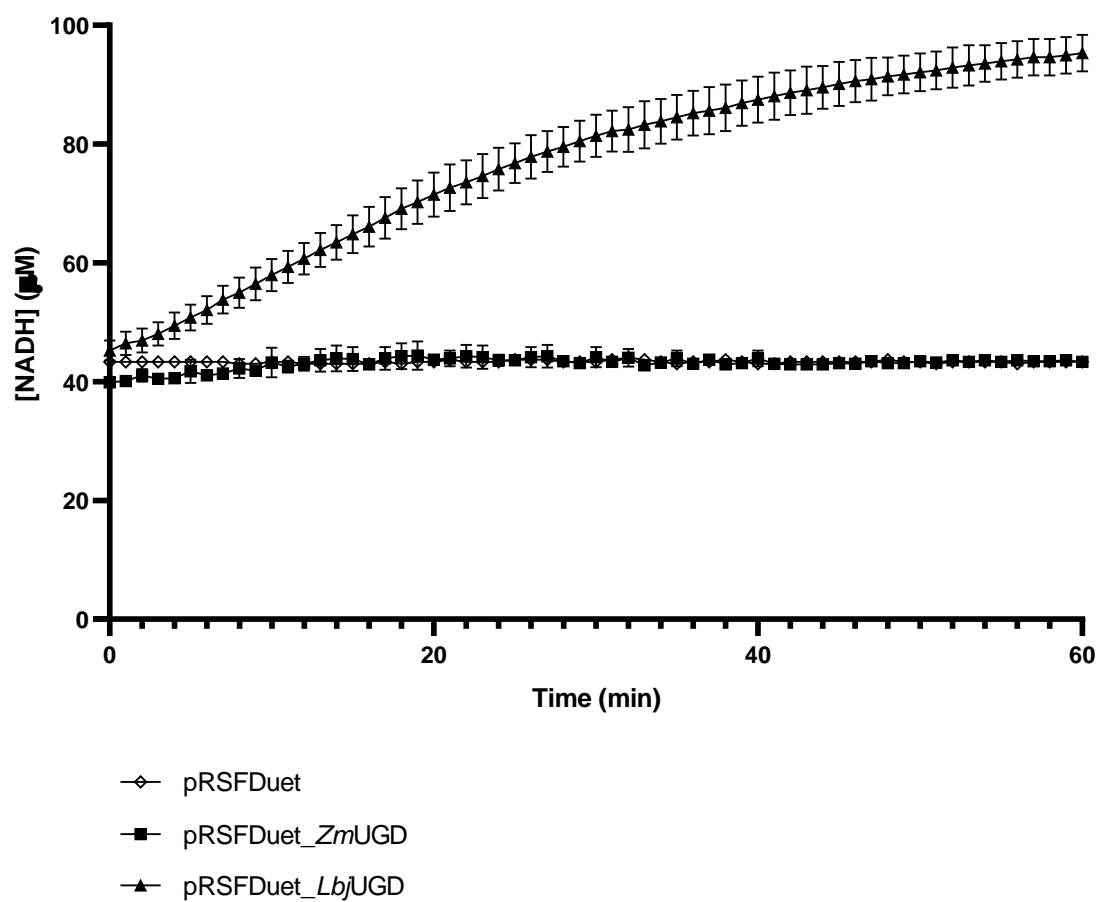


Figure S2. *In vitro* reduced nicotinamide adenine dinucleotide (NADH) production over time in engineered *Escherichia coli* at pH 7.2 as obtained by monitoring absorbance at 340 nm.

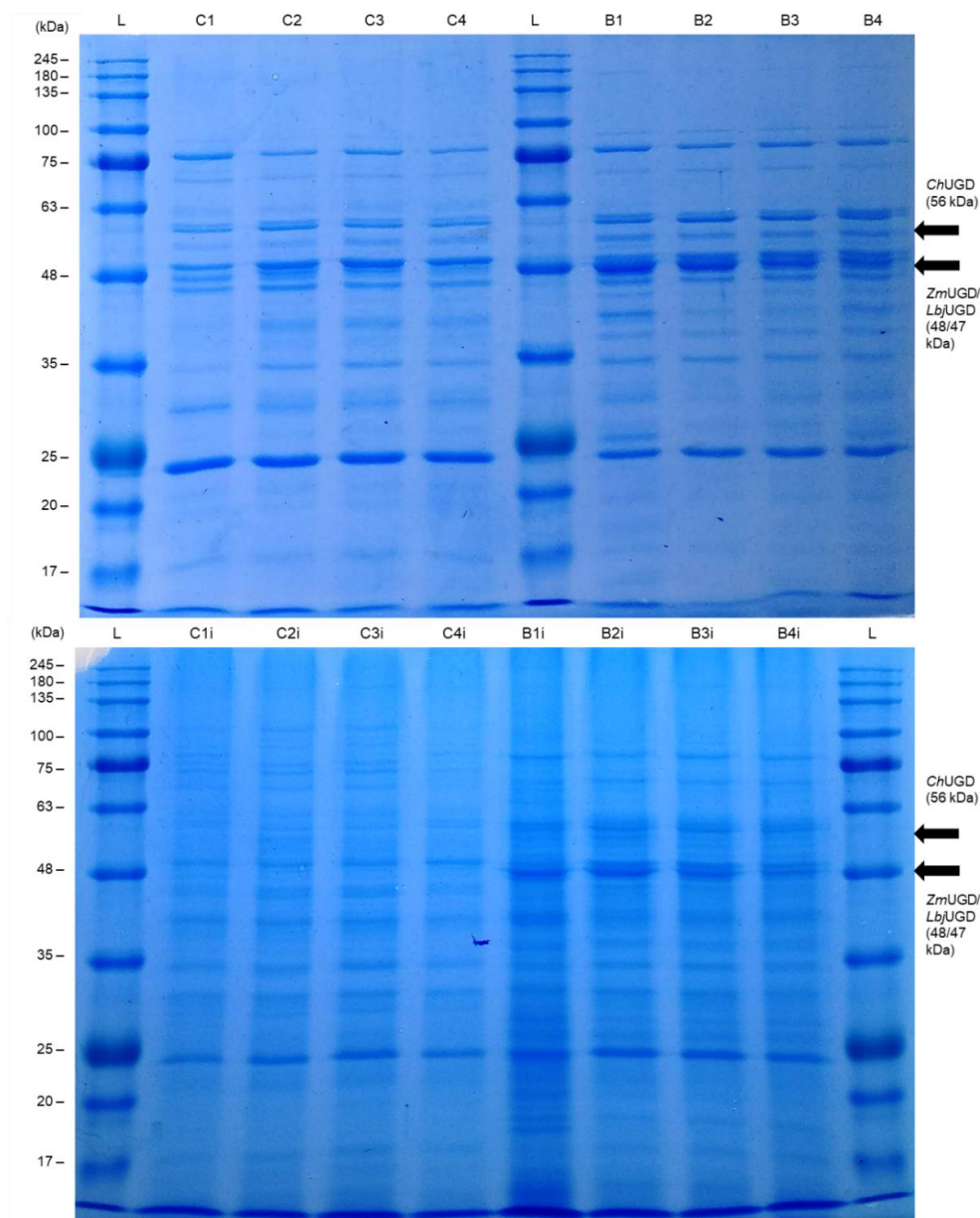


Figure S3. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels from *Saccharomyces cerevisiae* cell extracts (15 µg). In the upper gel, the samples loaded were: cell extract from *S. cerevisiae* CEN.PK2-1C wild-type (C1), cell extract from *S. cerevisiae* CEN.PK2-1C expressing *ChUGD* (C2) cell extract from *S. cerevisiae* CEN.PK2-1C expressing *ZmUGD* (C3), cell extract from *S. cerevisiae* CEN.PK2-1C expressing *LbjUGD* (C4), *S. cerevisiae* BY4741 wild-type (B1), *S. cerevisiae* BY4741 expressing *ChUGD* (B2), *S. cerevisiae* BY4741 expressing *ZmUGD* (B3) and *S. cerevisiae* BY4741 expressing *LbjUGD* (B4). The lower gel shows the proteins in the corresponding insoluble phases. L: NZYColour Protein Marker II ladder.

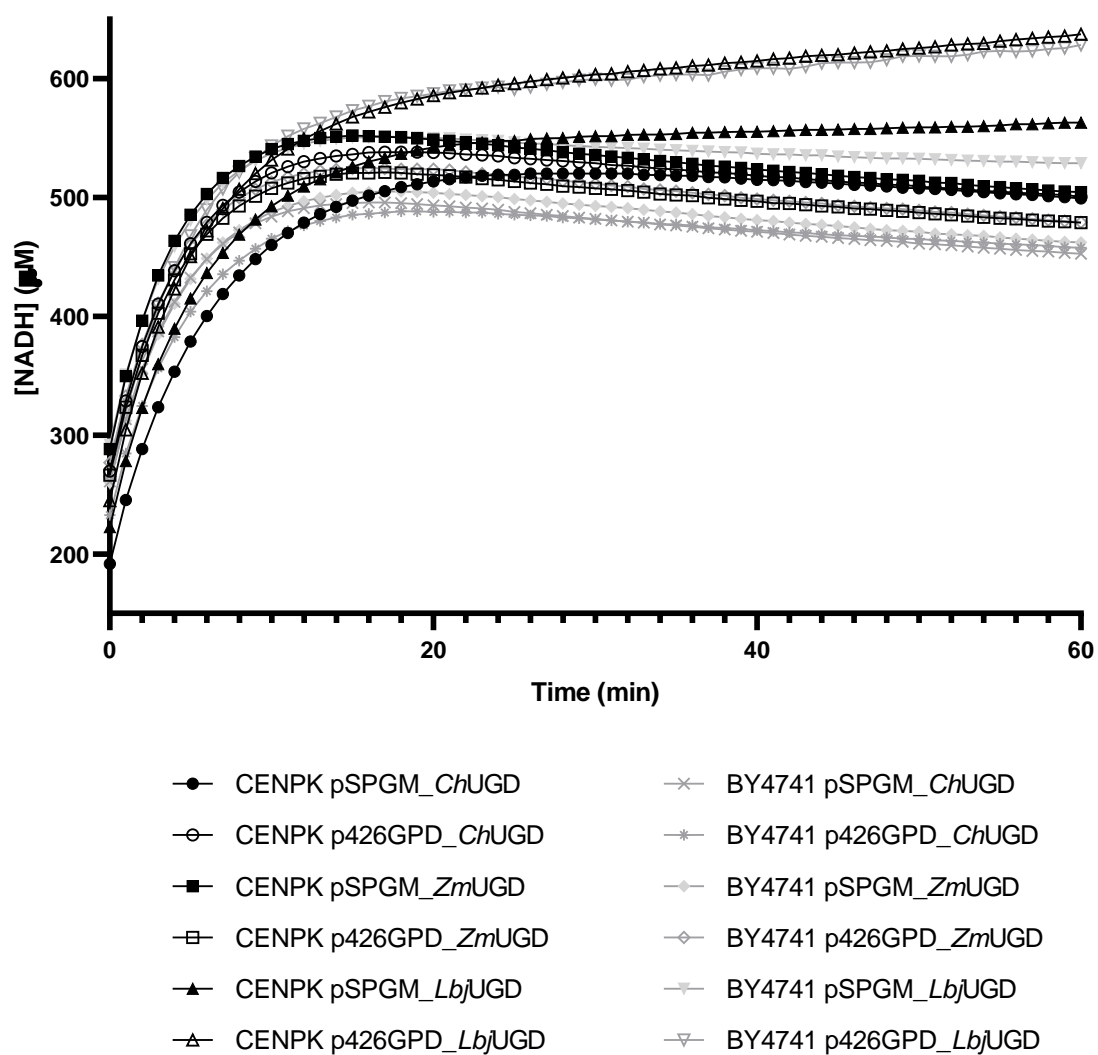


Figure S4. *In vitro* reduced nicotinamide adenine dinucleotide (NADH) production over time in engineered *S. cerevisiae* yeasts at pH 9.2 obtained by monitoring absorbance at 340 nm.