Electronic Supplementary Material (ESI) for RSC Advances. This journal is © The Royal Society of Chemistry 2017

	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
Debaryomyces nepalensis	MSIKLNSGY EMPLVGFGC <mark>W</mark> KVDNATCADT VYNAIKVGYR LF <mark>D</mark> AAM <mark>DY</mark> GN
scheffersomyces shehatae	MSPSP IPAFKLNNGL EMPSIGFGCW KLDKSTAADQ VYNAIKAGYR LFDGAEDYGN
Candida tenius Spathaspora passalidarum	MSAS IPDIKLSSGH LMPSIGFGCW KLANATAGEQ VYQAIKAGYR LFDGAEDYGN MATIKLSSGH LMPLVGFGCW KVDNATAADQ IYNAIKAGYR LFDGAEDYGN
Debaryomyces hensenii	
Pichia pastoris	M ATLLKLNNGL KLPQVGLGV <mark>W</mark> KIPNELTAET VYNAIKQGYR LF <mark>D</mark> GAE <mark>DY</mark> GN
Talaromyces emersonii	MA TPTIKLNSGY DMPLVGFGLW KVNRETCADQ VYEAIKAGYR LFDGACDYGN
Clustal Consensus	:*** :* :*:* * *: :*:*** *** ********
	70 80 90 100 110 120
Debaryomyces nepalensis	CKEIGEGINR ALDEGLVARD ELFITSKLWN SYHDPKNVEL ALKKVLSDMK LDYIDLFLIH EQEVGEGVKR AIDEGIVTRE EIFLTSKLWN NYHDPKNVET ALNKTLKDLK VDYVDLFLIH
scheffersomyces shehatae Candida tenius	EQEVGEGVRR AIDEGIVTRE EIFLTSKLWN NIHDPRNVET ALNKTLKDLK VDIVDLFLIH EKEVGDGVRR AIDEGLVKRE EIFLTS <mark>K</mark> LWN NYHDPRNVET ALNKTLADLK VDIVDLFLIH
Spathaspora passalidarum	EKEVGDGLKR AIDEGLVKRE ELFITSKLWN NYHDPKNVET ALNRTLSDLQ LDYVDLFLIH
Debaryomyces hensenii	CKEIGEGINK ALDEGLVARD ELFITS <mark>K</mark> LWN SYHDPKNVEL ALKKVLSDMK LDYLDLFLIH
Pichia pastoris	EKEVGQGVRR AIDEGLVKRE DLFIVS <mark>K</mark> LWN NYHHPDNVGK ALDRTLSDLG LDYLDLFYIH
Talaromyces emersonii	EVEAGQGVAR AIKEGIVKRE DLFIVS <mark>K</mark> LWN TFHEADKVEP IARKQLADWG LDYFDLYLIH
Clustal Consensus	* *:*: : *:.**:* *: ::*:.**** .:*:* : * * :**.**: **
	130 140 150 160 170 180
Debaryomyces nepalensis	FPIAFKFVPF EEKYPPAFYC GDGDN-F HYEDVPLLET WKAMEKLTKG GKAKSIGISN
scheffersomyces shehatae Candida tenius	FPIAFKFVPI EEKYPPGFYC GDGDN-F VYEDVPILET WKALEKLVKA GKIRSIGVSN FPIAFKFVPI EEKYPPGFYC GDGNN-F VYEDVPILET WKALEKLVAA GKIKSIGVSN
Spathaspora passalidarum	FPIAFKFVFI EEKIFFGFIC GDGNN-F VIEDVFILEI WAALEKLVAA GRINSIGVSN FPIAFKFVPL EEKIFPGFIC GDGNN-F HYENVPLLDT WKALEKLVQA GRIKSIGISN
Debaryomyces hensenii	FPIAFKFVPI EERYPPGFYC GDGDK-F HYENVPLADT WKAMEKLTKS GKVKSIGISN
Pichia pastoris	FPIAFKFVPL EEKYPPAFYC GDGNN-F HYEDVPLLDT YRALERLVDA GRIKSLGVSN
Talaromyces emersonii	FPIALKYVDP AEIYPPG-WT GTKKEV EFSNATIQET WQAMETLVDK KLTRSIGISN
Clustal Consensus	****:*:* * ***.: * .: . :.:.: :* ::*:* *. :*:*:**
	190 200 210 220 230 2 40
Debaryomyces nepalensis	FSAALIYDLL RGAEIKPAVL QIEHHPYLQQ PRLIEYVQSQ GIAITAYSSF GPQSFL
scheffersomyces shehatae Candida tenius	FPGALLLDLF RGATIKPAVL QVEHHPYLQQ PKLIEYAQKV GITVTAYSSF GPQSFVIMNQ FPGALLLDLL RGATIKPAVL QVEHHPYLQQ PKLIEFAQKA GVTITAYSSF GPQSFVIMNQ
Spathaspora passalidarum	FPGALIYDLV RGATIKPAVL QIEHHPYLQQ PKLIEYVQKQ GIAITAYSSF GPQSFLBLNQ
Debaryomyces hensenii	FSGALIYDLL RSAEIKPAVL QIEHHPYLQQ PRLVEYVQSQ NIAITGYSSF GPQSFLELKH
Pichia pastoris	FNGALLQDLL RGARIKPVAL QIEHHPYLVQ QKLIEYAQSE DIVVVAYSSF GPQSFL <mark>B</mark> LKV
Talaromyces emersonii	FSAQLIMDLL RYARIRPATL QIEHHPYLTQ QALVEYVQKE GIAVTAYSSF GPLSFLELGH
Clustal Consensus	* . *: **. * * *:** *:***** * *:*:.*:**** ** **:*:
	250 260 270 280 <u>29</u> 0 300
Debaryomyces nepalensis	SKALDTPTLF EHKTITSIAD KYKKTPAQVL LRWASQRDIA IIPKSNNPDR LLQNLEVND-
scheffersomyces shehatae Candida tenius	GRALNTPTLF EHDVIKAIAA KHNKVPAEVL LRWSAQRGIA VIP <mark>KSNLPER</mark> LVQNRSFND- GRALNTPTLF AHDTIKAIAA KYNKTPAEVL LRWAAQRGIA VIP <mark>KSNLPER</mark> LVQNRSFNT-
Spathaspora passalidarum	NRALNIFILF ANDIIKAIAA KINKIFALVI LKWARQKGIA VIFANALFEN LVQKSFNI- NRALNIFILF EHDIIKSIST RLNKTPAQVL LRWATQRNIA VIPKSNNPAR LAQNLDVTS-
Debaryomyces hensenii	SKALDTPTLF EHKTIKSIAN KNKKTPAQVL LRWASQRNIA VIPKSNNPDR LLQNLEVND- NKALTAVSLF EHDVIKKIAQ AHNRSAGEVL LRWATQRGLA IIPK <mark>SS</mark> KPER LSSNLHINS-
Pichia pastoris	NKALTAVSLF EHDVIKKIAQ AHNRSAGEVL LRWATQRGLA IIPASSKPER LSSNLHINS-
Talaromyces emersonii	QVAKDTPLLF EHSTVKSIAE KHGKTPAQVL LRWATQRNIA VIP <mark>KSN</mark> NPG <mark>R</mark> LAQNLDVTA-
	NKALTAVSLF EHDVIKKIAQ AHNRSAGEVL LRWATQRGLA IIPASSKPER LSSNLHINS- QVAKDTPLLF EHSTVKSIAE KHGKTPAQVL LRWATQRNIA VIPASNNPGR LAQNLDVTA- * : ** * *: :** ***::**.:* :**** ** * * *
Talaromyces emersonii	QVAKDTPLLF EHSTVKSIAE KHGKTPAQVL LRWATQRNIA VIPKSNNPGR LAQNLDVTA- * : ** * *: :** ***::**.:* :****. * *
Talaromyces emersonii Clustal Consensus	QVARDTPLLF EHSTVKSIAE KHGKTPAQVL LRWATQRNIA VIP LAQNLDVTA- * : **: : <td:< td=""> : : <td< td=""></td<></td:<>
Talaromyces emersonii Clustal Consensus Debaryomyces nepalensis	QVAKDTPLLF EHSTVKSIAE KHGKTPAQVL LRWATQRNIA VIPKSNNGG LAQNLDVTA- * : ** * *: :** ***::**.** :**** . 310 320 330 FNLSKEDFDE ISKLDQDLRF NPWDWDTKN RIPIFA
Talaromyces emersonii Clustal Consensus Debaryomyces nepalensis scheffersomyces shehatae	QVAKDTPLLF EHSTVKSIAE KHGKTPAQVL LRWATQRNIA VIPKSNNGR LAQNLDVTA- * : ** * *: :** ***::**.** :***** * : ** *
Talaromyces emersonii Clustal Consensus Debaryomyces nepalensis scheffersomyces shehatae Candida tenius	QVARDTPLLF EHSTVKSIAE KHGKTPAQVL LRWATQRNIA VIPASNAPGE LAQNLDVTA- * : ** * *: :** ***::**.:* :**** * * *
Talaromyces emersonii Clustal Consensus Debaryomyces nepalensis scheffersomyces shehatae	QVAKDTPLLF EHSTVKSIAE KHGKTPAQVL LRWATQRNIA VIPKSNNGR LAQNLDVTA- * : ** * *: :** ***::**.** :***** * : ** *
Talaromyces emersonii Clustal Consensus Debaryomyces nepalensis scheffersomyces shehatae Candida tenius Spathaspora passalidarum	QVAKDTPLLF EHSTVKSIAE KHGKTPAQVL LRWATQRNIA VIFKSNNGG LAQNLDVTA- * : ** * *: :** ***::**.** :**** * * *
Talaromyces emersonii Clustal Consensus Debaryomyces nepalensis scheffersomyces shehatae Candida tenius Spathaspora passalidarum Debaryomyces hensenii	QVARDTPLLF EHSTVKSIAE KHGKTPAQVL LRWATQRNIA VIPASIANDGE LAQNLDVTA- * : ** * *: :** ***::*** :*****. * * * . 310 320 330 FNLSKEDFDE ISKLDQDLRF NPWDWDTKN RIPIFA FELTKEDFEE ISKLDIGLRF NDPWDWDNIPIFV FDLTEEDFAA ISALDINLRF NDPWDWDNIPIFV FDLTEEDFAA ISALDINLRF NDPWDWDNIPIFV FDLSKEDFEE ISKLDQELRF NDPWDWDKIPIFA

Fig. S1. Multiple alignments of DnAR and six other aldose (xylose) reductases. Sequences were aligned using ClustalW and were rendered using BioEdit 7.0.0. The residues involving in substrate binding (green), catalysis (red) as well as the important residues for cosubstrate binding (black) are highlighted. The clustal consensus is given below the alignments. Asterisks indicate high conservation of residues, whereas double and single dots represent decreasing conservation consecutively. Underlined regions indicate determined peptide sequences of DnAR by peptide mass fingerprinting using LC-MS.

Mass spectroscopy (LC-MS) analysis of purified DnAR

The pH of the purified protein sample was adjusted to ~8.5 using 50 mM ammonium carbonate. To this sample (2 ml), ice cold trypsin (0.2 U/µl) was added in 1:30 (enzyme:protein) ratio and incubated at 55 °C for 2 h. The digested peptides were vacuum dried and reconstituted in 15 µl of 2% (v/v) acetonitrile with 0.1% (v/v) formic acid. The digested peptides were subjected to 70 min reversed phase chromatography (RPC) gradient, followed by acquisition of the data on linear trap quadropole (LTQ)-Orbitrap-Mass spectroscopy (Thermo Scientific, USA). The generated peptide fragments data were searched against Uniprot Swiss-Prot database, UniprotTrEMBL and yeast RefSeq database (from NCBI). Mass spectroscopic analysis of the purified protein was performed at the Centre for Cellular and Molecular Platforms (C-CAMP), Department of Biotechnology, Government of India.