The EMBO Journal vol.3 no.3 pp.611-616, 1984

Molecular cloning, DNA structure and expression of the *Escherichia coli* D-xylose isomerase

Kathleen A. Briggs[1]*, William E. Lancashire and Brian S. Hartley

Centre for Biotechnology, Biochemistry Building, Imperial College of Science and Technology, London, SW7 2AZ, UK

[1]Present address: GENETICA, 160 quai de Polangis, 94340 Joinville-le-Pont, Paris, France
*To whom reprint requests should be sent
Communicated by B.S. Hartley

The D-xylose isomerase (EC 5.3.1.5) gene from *Escherichia coli* was cloned and isolated by complementation of an isomerase-deficient *E. coli* strain. The insert containing the gene was restriction mapped and further subcloning located the gene in a 1.6-kb *Bgl*II fragment. This fragment was sequenced by the chain termination method, and showed the gene to be 1002 bp in size. The *Bgl*II fragment was cloned into a yeast expression vector utilising the *CYC1* yeast promoter. This construct allowed expression in *E. coli* grown on xylose but not glucose suggesting that the yeast promoter is responding to the *E. coli* catabolite repression system. No expression was detected in yeast from this construct and this is discussed in terms of the upstream region in the *E. coli* insert with suggestions of how improved constructs may permit achievement of the goal of a xylose-fermenting yeast.
Key words: E. coli/expression/sequence/xylose isomerase/yeast

Introduction

Hydrolysed hemicellulose components of wood and agricultural wastes have great potential as feedstocks for the production of chemicals, particularly ethanol, by microbial processes. These pentose sugars, mainly xylose, are abundant in a variety of common lignocellulose residues and can be readily recovered by mild acid hydrolysis. Fermenting yeasts, for example, *Saccharomyces cerevisiae* would be ideal ethanol-tolerant organisms for converting D-xylose to ethanol, but this yeast is unable to carry out this anaerobic fermentation. *S. cerevisiae* can, however, utilise the pentulose D-xylulose anaerobically for both growth and fermentative ethanol production (Wang *et al.*, 1980) and indeed two-stage processes have been developed using an immobilised xylose isomerase (EC 5.3.1.5) to carry out the xylose-xylulose interconversion (Gong *et al.*, 1981).

The present work forms part of a program to genetically engineer a yeast, using recombinant DNA techniques, which will contain an *in situ* xylose isomerase activity and therefore should ferment xylose directly. The aspect described here is the isolation by cloning of the xylose isomerase gene from *E. coli* including subcloning, characterisation and sequence determination. The data are analysed for possible pointers to the organisation and regulation of the xylose operon in *Escherichia coli* in comparison to that already known for *Salmonella typhimurium* (Shamanna and Sanderson, 1979a, 1979b).

Results

Cloning of D-xylose isomerase

Our strategy for cloning was to complement a xylose isomerase-deficient mutant of *E. coli*. Strains of *E. coli* were screened enzymatically for isomerase activity and also xylulokinase and xylose permease activities. A suitable recipient was found to be JA221/xyl$^-$.

We noted that four plasmids in the Clarke-Carbon ColE1 gene library had been identified as containing inserts that complement undefined xylose mutations (Clarke and Carbon, 1976). Plasmid DNA was isolated from these clones and the sizes were estimated by agarose gel electrophoresis as ~30 kb (pLC1-3), 26 kb (pLC10-15) and 17 kb (pLC32-9). The plasmids were all capable of transforming strain JA221/xyl$^-$ (isomerase-deficient) and strain R8 (kinase-deficient) to xylose-positive, indicating the presence of both DXI and DXK genes and thus probably a substantial portion of a putative xylose operon.

Restriction mapping of the clones

To plan a subcloning and a sequencing strategy for the isomerase gene, we restriction mapped the three plasmids using single, double and triple digestions with various restriction enzymes. Figure 1 shows the pattern deduced. As expected, it is clear that all three plasmids have overlapping sequences in which the gene must be contained.

Subcloning of the xylose isomerase gene

To extract the gene from the ColE1 plasmid we cloned the 12.5-kb *Hind*III fragment from pLC10-15 into vector pJDB207. The resulting hybrid plasmid pKB3 (Figure 2) was still capable of transforming both the isomerase-deficient and kinase-deficient mutants. To locate the isomerase gene more precisely we performed further subcloning experiments (see Figure 2). Plasmid pKB3 was partially digested with *Bgl*II, religated and re-transformed into selection strains. A similar treatment was also carried out with *Sal*I. Figure 2 shows the deletion plasmids (i) to (vi) together with their complementing capabilities. It is clear that the 1.6-kb *Bgl*II fragment contains the whole of the coding region for xylose isomerase and that the *Sal*I site is included in the gene. The same fragment would also appear to contain part of the xylulokinase coding region.

As an initial attempt to express the xylose isomerase gene in yeast, subclonings were made in the yeast expression vector pEX-2, constructed by J. Ernst, (J. Ernst and F. Sherman, unpublished results). This makes use of the *CYC1* promoter and terminator regions with a suitable multiple cloning site region in which inserts may be made to allow expression from the *CYC1* promoter. The 1.6-kb *Bgl*II fragment from pKB3 was inserted into the *Bam*HI site in pEX-2 and cloned by transformation and ampicillin selection in *E. coli* (JA221) followed by identification of hybrids by mini-plasmid preparations. When these were further analysed by EMB-xylose plating and restriction mapping, we were interested to find that when the insert was in the correct orientation,

© IRL Press Limited, Oxford, England.

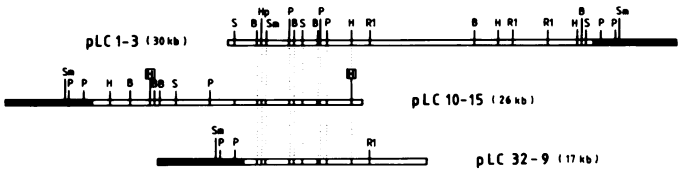

Fig. 1. Restriction maps of pLC1-3, pLC10-15 and pLC32-9. The ColE1 DNA (6.3 kb) is indicated by the hatched regions and the segments corresponding to the poly(dT.dA) junctions shown as solid shading. The *Hind*III sites enclosed by boxes ([H]) indicate the region of pLC10-15 cloned into pJDB207 to construct plasmid pKB3. Abbreviations of restriction enzymes are: B = *Bgl*II, R1 = *Eco*RI, Hp = *Hpa*I, H = *Hind*III, P = *Pst*I, S = *Sal*I, Sm = *Sma*I.

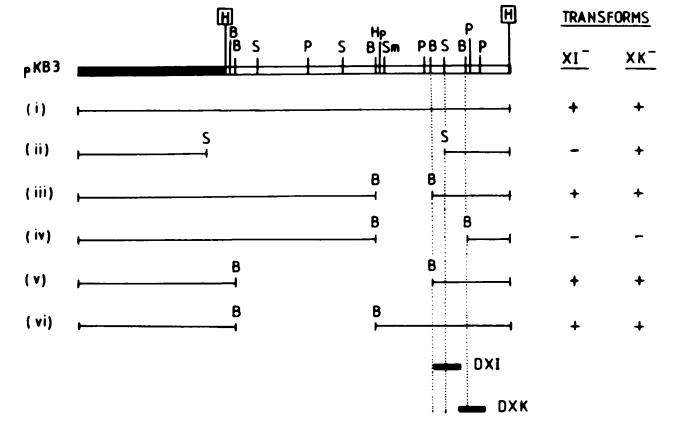

Fig. 2. Subcloning of pKB3. The linear representations (**i**–**vi**) of the plasmids show the sections of DNA absent in the subclones. Plasmid (**i**) is intact pKB3, (**ii**)–(**vi**) are subclones. Transformation results in complementation (+) or non-complementation (−) of DXI or DXK mutations. Abbreviations H = *Hind*III, B = *Bgl*II, S = *Sal*I, P = *Pst*I, Hp = *Hpa*I, Sm = *Sma*I, XI⁻ = *E. coli* xylose isomerase-minus mutant JA221/xyl⁻, XK⁻ = *E. coli* xylulokinase-minus mutant R8. The vector DNA is shown as solid shading in the diagram. The tentative positions of the D-xylose isomerase gene (DXI) and the D-xylulokinase gene are indicated.

transcription and translation was apparently occurring from the yeast promoter. Guarante and Ptashne (1981) have also demonstrated expression in *E. coli* from the yeast *CYC1* promoter. Hybrid pKB16X (Figure 3) is an example of this construct.

Expression of xylose isomerase activities in E. coli and yeast subclones

The *E. coli* strain JA221/xyl⁻ containing the subclone pKB16X described above was grown on synthetic medium using either glucose or xylose as carbon source. Xylose isomerase activities were measured as described in Materials and methods. Very little activity (0.03 μmol D-xylose isomerised/mg protein/min) was detected in glucose grown cells whereas 0.84 units were measured in xylose grown extracts. This unexpected finding is discussed later.

No xylose isomerase activity could be detected in yeast cells transformed with pKB3 in which the 12.5-kb *Hind*III insert of *E. coli* DNA contains the whole of the *xylB-xylA* region but no yeast promoter. This suggests, not surprisingly, that yeasts do not recognize the *E. coli* transcription and/or translation signals for this operon.

No xylose isomerase activity could be detected in yeast cells transformed with pKB16X. This lack of expression in yeast is probably not due to instability of the xylose isomerase within the yeast cell. Incubation of an *E. coli* (JA221/xyl⁺) containing DXI activity with an extract of the yeast used in transformations (DBY 747) for 10 min at 30°C produced no decrease in specific activity of this enzyme. Since this yeast extract is rich in proteases, the D-xylose isomerase appears to be a quite stable enzyme.

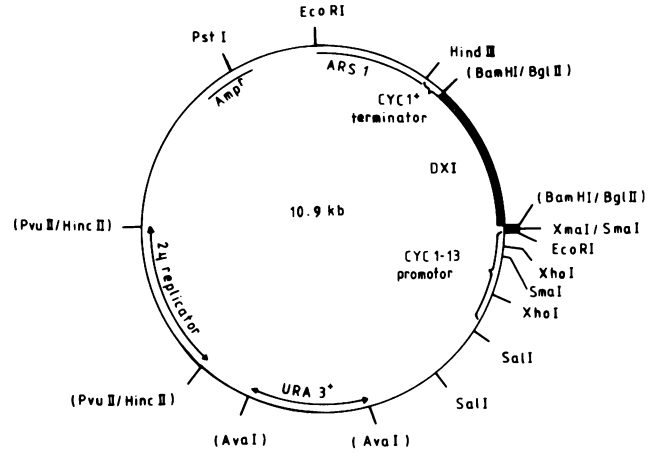

Fig. 3. Restriction map of pKB16X. The *Bgl*II fragment (1.6 kb) from plasmid pKB3 was ligated into the *Bam*HI site in yeast expression vector pEX-2 as described in Materials and methods.

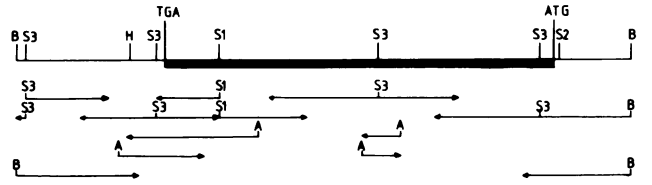

Fig. 4. Sequencing strategy of the 1.6-kb *Bgl*II fragment containing the xylose isomerase gene (DXI). The M13 clones are aligned with the restriction map and the arrowheads indicate the extent of DNA sequence read. Restriction site abbreviations are as follows: A = *Alu*I, B = *Bgl*II, H = *Hae*II, Sl = *Sal*I, S3 = *Sau*3A, S2 = *Sst*II.

DNA sequencing of D-xylose isomerase

The above subcloning experiments indicated the presence of the whole of the DXI gene and part of the DXK gene in the 1.6-kb *Bgl*II fragment of pKB3. This section of DNA was therefore chosen for nucleotide sequence determination.

The fragment was isolated from a complete *Bgl*II digest of pKB3 by the low-melting point agarose method. The DNA was extracted with phenol, extracted with ether and precipitated with ethanol prior to further use. The purified fragment was digested with restriction enzymes *Sau*3A, *Sal*I and *Alu*I, ligated into appropriately digested M13 vectors mp8 and mp9 and transformed into JM101. The sequencing strategy is shown in Figure 4 and was deduced from clone cross-hybridisation and overlap analysis. Sequence data were loaded onto a Cyber174 computer using Fortran IV programs. Overlapping sequences were initially identified by eye and then, by means of the Staden program SQRVCM, the entire 1.6-kb nucleotide sequence was calculated in both directions, i.e., both DNA strands. The sequence of the *Bgl*II fragment is given in Figure 5.

Analysis of the nucleotide sequence

The *Bgl*II fragment was determined from the sequence data to be 1616 bp in length. The Staden computer program TRNTRP3 was used to translate the DNA sequence into the three letter amino acid code in both directions and in each of

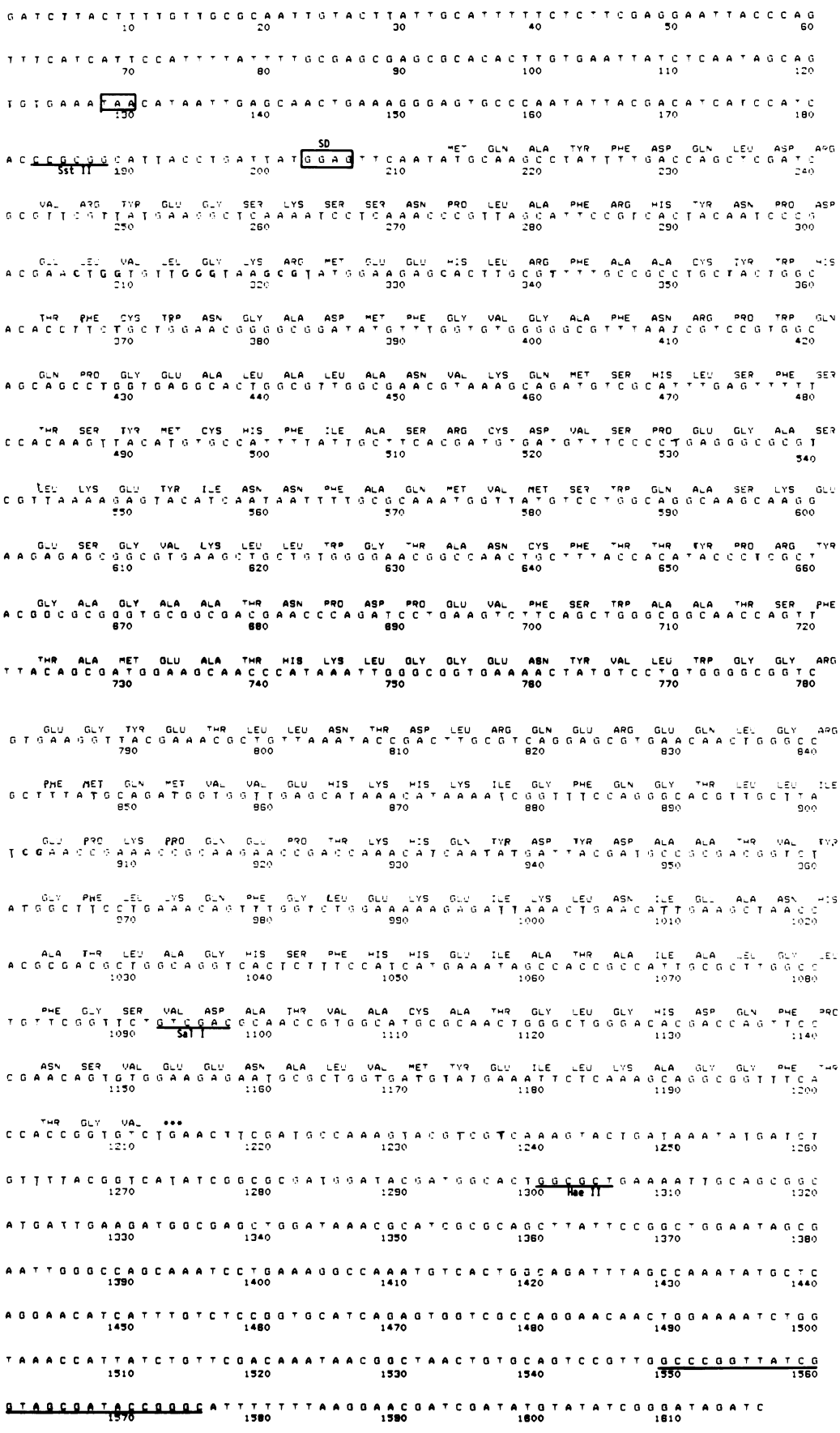

Fig. 5. Nucleotide sequence of the 1.6-kb *Bgl*II fragment showing the predicted amino acid sequence of D-xylose isomerase. The ribosome-binding site for the isomerase gene is shown boxed. Restriction sites for *Hae*II, *Sal*I and *Sst*II are marked. The xylulokinase termination codon at position 128 is boxed. The potential hairpin loop sequence is shown underlined from positions 1549 to 1575.

the three possible reading frames. Only one direction and one reading frame contains a significant stretch of open reading-frame, from position 126–1212, which is compatible with the DXI gene.

Putative initiation codons (ATG) occur at positions 201 and 213, so to decide which is the most likely, the sequence 6–10 bp downstream was examined for evidence of a ribosome-binding site (Shine and Dalgarno, 1974, 1975). There is no such sequence corresponding to MET 201 whereas the sequence d(A-T-G-G-A-G) exists between −12 and −7 from MET 213 as shown in Figure 5. It is predicted therefore, in the absence of N-terminal protein data, that the DXI gene initiates at position 213 ATG and terminates at 1212.

The subcloning data on plasmid pKB3 indicated that part of the DXK gene was also encoded in the 1.6-kb *Bgl*II fragment. Examination of the sequence in all three reading frames showed the longest stretch of open reading frame (apart from the DXI coding region) to terminate at position 128, 85 bp from the DXI initiation codon. This almost certainly corresponds to the carboxy terminus of the DXK gene.

Termination of transcription

In the simplest case of termination of transcription in the absence of *rho* or any other factors, several common features have been observed (Platt, 1981; Rosenberg and Court, 1979). Formation of a hairpin structure in the transcript as a result of a GC-rich inverted repeat causes RNA polymerase to pause while effective termination requires a suitable number of consecutive U residues following the hairpin.

The sequence of the DXI gene elucidated here predicts termination of translation at the TGA in position 1212. Although this site is not closely followed by the requisite structures described above, computer analysis using Staden program HAIRPN reveals a GC-rich hairpin followed by seven T residues, with several stop codons, situated ~360 bp upstream of the position 1212 TGA and stretching from positions 1549 to 1575. Other energetically favourable hairpin structures occur at positions 1302–1319 and 1362–1377 and have free energies of −6.1 kcal and −8.0 kcal respectively, calculated according to Tinoco *et al.* (1973). The larger hairpin from 1549 to 1575 has a free energy of −27.2 kcal making it considerably more stable than the other two. It is likely that the last gene in the xylose operon is the isomerase gene, and that the DNA sequence downstream of the gene, including the hairpin loops, constitute a transcription termination site for the entire operon. This hypothesis is supported by the fact that there appear to be only two potential hairpins, neither of which is particularly stable (free energies of −5.2 kcal and −6.4 kcal) towards the end of the DXK gene or between the DXK and DXI genes. Hence only limited intra-operon termination of transcription may be permitted.

Codon usage

The codon usage for the DXI gene has been analysed from Figure 5. It is clear that the situation is non-random and is quite typical of many *E. coli* genes. Codon usage has been suggested as being a factor modulating translation efficiency (Fiers *et al.*, 1976). Genes coding for abundant proteins use the majority of preferred codons. In the case of less abundant proteins more non-preferred codons may be employed. Proteins present in moderate amounts, most probably including DXI utilise a combination of the various codons. The correlation between tRNA levels and codon usage is discussed by Ikemura (1981).

Discussion

The D-xylose operon of E. coli

In *S. typhimurum* four genes encode inducible xylose metabolism: in order, *xylT-xylR-xylB-xylA* (Shamanna and Sanderson, 1979b). The *xylR* gene produces an activator protein which binds xylose and then binds to an operator/promoter for a xylose transport gene(s) (*xylT*) and to another operator/promoter for the *xylB* (D-xylulokinase)-*xylA* (D-xylose isomerase) genes.

The genes in *E. coli* seem to be similar, since xylose transport, D-xylulokinase and D-xylose isomerase are under positive regulatory control (David and Weismeyer, 1970) and the genes map at 80 minutes in *E. coli* (78 minutes in *S. typhimurium*) adjacent to a *glyS* gene. D-Xylulokinase is induced before D-xylose isomerase in both species, so the promoter for these genes is assumed to be adjacent to *xylB* (Shamanna and Sanderson, 1979a). Our studies confirm this arrangement, since pLC 10-15 contains *xylA, xylB* and *glyS* and subcloning and DNA sequencing prove the gene order. Figure 6 shows the tentative structure within the *Hind*III fragment of pLC 10-15. The *xylA* gene is 1002 bp in length separated by 82 bp from the C terminus of *xylB*. If we assume *xylB* to be ~1.5 kb by analogy with the D-xylulokinase gene of *Klebsiella aerogenes* (Neuberger and Hartley, 1979), the 900 bp 5′ to *xylB* must contain the operator/promoter and also part or all of *xylR*. The plasmid-borne genes are xylose-inducible in agreement with this hypothesis.

Transcription and translation of the D-xylose isomerase operon

Transcription must begin in the 900-bp region 5′ to *xylB* and terminate in the observed hairpin loops 3′ to *xylA*. There are two ATG codons in the open reading frame of *xylA* (201 and 213 in Figure 5) but only the latter is preceded by a plausible Shine-Dalgarno sequence, GGAG, −10 to −7 relative to ATG-213. Hence we assume that the amino acid sequence of D-xylose isomerase is as shown in Figure 5.

Expression of D-xylose isomerase in E. coli and yeast

Xylose isomerase activity was found to be xylose-inducible, as expected, when plasmid pKB3 containing the whole of the *xylB-xylA* region was transformed into *E. coli*. This inducible effect was also observed when the xylose isomerase gene was introduced into *E. coli* in plasmid pKB16X which contains the *xylA* gene inserted 212 bp from the yeast mutated initiation codon (ATA) of the yeast gene encoding iso-1-cytochrome *c* (*cycl-13*, see Sherman and Stewart, 1983). This construct contains no *E. coli* transcription signals, so it is probable that transcription is from the yeast *CYC1* promoter, followed by translation from the ribosome-binding site upstream of the XI gene. This interpretation is strengthened by the observation that a plasmid pVK1X, containing the

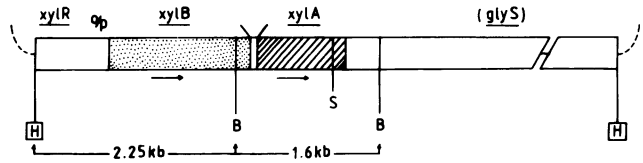

Fig. 6. Arrangement of the *xyl* operon in pKB3. A linear representation of pKB3 showing the positions of the genes and the tentative position of the operator/promoter site (O/P). The arrows indicate the directions of transcription. Abbreviations are as follows: H = *Hind*III boundaries between insert and pJDB207 DNA, B = *Bgl*II, S = *Sal*I.

same insert in the opposite orientation, expressed no xylose isomerase activity (V. Kumar, personal communication). But why, if this is the case, should the xylose isomerase be apparently xylose-inducible? It is unlikely that the yeast *CYC1* promoter would be activated by the *E. coli* xylose-activator protein. A more plausible explanation is that this promoter responds to the *E. coli* catabolite repression system. Glucose produces strong catabolite repression and xylose is probably much weaker in this respect.

In contrast to *E. coli*, no measurable xylose isomerase activity and presumably no expression of the gene occurred in yeast transformed with the pKB16X construct. There are many possible explanations for this, but we believe the most likely to be that yeast ribosomes do not recognise the *E. coli* translation signals immediately upstream of the xylose isomerase coding region and that translation is also not occurring from the *CYC1* promoter region. We have demonstrated that lack of expression of this construct in yeast is probably not due to instability of the xylose isomerase within the yeast cell. It is more likely that lack of expression is due to the 212-bp distance between the normal initiation point for cytochrome *c* and the initiation codon for xylose isomerase.

In theory initiation should commence, after the *CYC1* promoter, at the first ATG in the insert. This would correspond to that which is situated 12 bp upstream of the true XI gene initiation codon. This would result in a small extension to the N-terminal end of the protein and result in conformation inactivation of the enzyme. This and other possible problems associated with the 212-bp sequence between the initiation codon and the promoter are under investigation and we are making more 'perfect' constructs and utilising other yeast promoters in order to achieve our goal of a xylose utilising yeast. While this paper was in preparation another group reported the isolation of a DNA fragment containing the xylose isomerase gene, but no DNA sequence details were given (Ho *et al.*, 1983).

Materials and methods

Strains and plasmids

The following strains are derivatives of *E. coli* K12. The D-xylose isomerase mutant JA221 (*rec*A1 *leu*B6 *trp*△E5 *hsd*M$^+$ *hsd*R$^-$ *lac*Y *xyl* C600) and the D-xylulokinase (EC 2.7.1.17) mutant R8 (F$^-$ *sup*E *xyl*K, W1485) were used as recipients for transformations. Strain AB4134 (F$^-$ *xyl gly*S3 metC-56 *lac*Y1 *gal*K2 *ara*-14 *thi*-1 *thi*5 *tsx*-57 *sup*E44) was used as a source of the marker *gly*S which flanks the *xyl* region. ColE1-*E. coli* hybrid plasmids pLC1-3, pLC10-15 and pLC32-9 which complement the *xyl* mutation (Clarke and Carbon, 1976) were obtained in strain JA200 (*rec*A1 *leu*B6 and *trp*△E5 *thr lac*Y) from B. Bachmann. The yeast-*E. coli* hybrid vector pJDB207 was obtained from J. Beggs. Yeast expression vector pEX-2 (J. Ernst and F. Sherman, unpublished results; Gritz and Davies, 1983) was obtained from J. Ernst *E. coli* strain JM101 (△*lacpro sup*E *thi* F′ *tra* D36 *pro*AB *lac*I^q Z △M15) was used as recipient in the transfections.

DNA isolation

Plasmid DNA was isolated from 200 ml cultures in L-broth medium after amplification with 150 μg/ml chloramphenicol and subsequently purified by standard procedures (Clewell, 1972). Small scale preparations for rapid screening were carried out by the procedure of Kado and Liu (1981).

Restriction endonuclease digestion and ligation of fragments

Enzymes were used according to suppliers' instructions. Digestion products were analysed by electrophoresis in submerged horizontal agarose gels (0.5%, 0.7% or 1%) in a Tris-borate buffer system.

The established procedures for ligation of fragments into vectors and transformation of *E. coli* strains were used (Cohen *et al.*, 1973). Xylose-plus phenotypes were identified by selection of transformants on Eosin methylene blue (EMB) indicator plates (Miller, 1972) containing 1% xylose and also containing 50 μg/ml ampicillin when appropriate. On this medium, positive colonies exhibit a very deep purple colouration, often with a metallic green sheen.

Cloning in M13

Replicative forms of phage M13 were prepared, restriction fragments inserted and hybrid phages identified as white plaques on 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) indicator plates using established procedures (Sanger *et al.*, 1980).

Nucleotide sequence analysis

M13 hybrid single-stranded DNA was prepared (Sanger *et al.*, 1980) and the templates used in the dideoxy chain-termination sequencing method of Sanger *et al.* (1977). The primer used was either a 17-nucleotide (Collaborative Research) synthetic primer or the 15-nucleotide primer (Biolabs). Templates were initially screened by 'T-tracking' to eliminate duplicates and identify potentially useful clones (Sanger *et al.*, 1980). The nucleotide sequence data obtained was compiled and analysed using the Staden computer programs (Staden, 1977, 1978, 1979).

Enzyme assays

Cultures of bacterial strains to be assayed for D-xylulokinase (DXK), D-xylose isomerase (DXI) or D-xylose permease (DXP) activity were grown in the minimal salts medium of Shamanna and Sanderson (1979). After harvesting by centrifugation and resuspension in 30 mM Tris-HCl pH 7.5, 1 mM dithiothreitol, cells were broken by sonication and then centrifuged at 15 000 *g* to obtain the crude cell-free extract.

Xylose isomerase activity was measured by the method described by Shamanna and Sanderson (1979) using the cysteine-carbazole colorimetric assay for detection of D-xylulose produced (Dische and Borenfreund, 1951).

D-Xylose permease was measured essentially as described by Shamanna and Sanderson (1979a).

Enzyme activity was measured in yeast samples prepared either by the toluene permeabilisation method (Maitra, 1970) or by the method of glass bead-breakage (Beggs *et al.*, 1980).

Yeast transformation

Sphaeroplasts were prepared and transformed with plasmid DNA by established procedures (Beggs, 1978).

Materials

Restriction endonucleases were purchased from Bethesda Research Laboratories, Inc., New England Biolabs and Boehringer Corporation Ltd. Phage T4 DNA ligase and M13 vectors mp8 and mp9 were obtained from New England Biolabs. Nucleoside triphosphates were obtained from P-L Biochemicals or New England Biolabs and 2-deoxyadenosine 5′-[α-^{32}P]triphosphate was purchased from Amersham International.

References

Beggs,J.D. (1978) *Nature,* **275**, 104-109.

Beggs,J.D. (1981) in von Wettstein,D., Friis,J., Keilland-Brandt,M. and Stenderup,A. (eds.), *Molecular Genetics in Yeast*, Munksgaard, Copenhagen.

Beggs,J.D., van den Berg,J., van Ooyen,A. and Weissman,C. (1980) *Nature,* **283**, 835-840.

Clarke,L. and Carbon,J. (1976) *Cell,* **9**, 91-99.

Clewell,D.B. (1972) *J. Bacteriol.,* **110**, 667-676.

Cohen,S.N., Chaong,A.C.Y., Boyer,H.W. and Helling,R.B. (1973) *Proc. Natl. Acad. Sci. USA,* **70**, 3240-3244.

David,J.D. and Weismeyer,H. (1970) *Biochim. Biophys. Acta,* **201**, 497-499.

Dische,Z. and Borenfreund,E. (1951) *J. Biol. Chem.,* **192**, 583-587.

Fiers,W., Contreras,R., Duerinck,F., Haegeman,G., Iserentant,D., Merregaert,J., MinJou,W., Molemans,F., Raeymaekers,A., Van den Berghe,A., Volckaert,G. and Ysebaert,M. (1976) *Nature,* **260**, 500-507.

Gritz,L. and Davies,J. (1983) *Gene,* **25**, 179-188.

Gong,C.-S., Chen,L.-F., Flickinger,M.C., Chiang,L.-C. and Tsao,G.T. (1981) *Appl. Environ. Microbiol.,* **41**, 430-436.

Guarante,L. and Ptashne,M. (1981) *Proc. Natl. Acad. Sci. USA,* **78**, 2199-2203.

Ho,N.W.Y., Rosenfeld,S., Stevis,P. and Tsao,G.T. (1983) *Enzyme Microb. Technol.,* **5**, 417-420.

Ikemura,T. (1981) *J. Mol. Biol.,* **146**, 1-21.

Kado,C.I. and Liu,S.T. (1981) *J. Bacteriol.,* **145**, 1365-1373.

Maitra,P.K. (1970) *J. Biol. Chem.,* **245**, 2423-2431.

Miller,J.H. (1972) *Experiments in Molecular Genetics*, published by Cold Spring Harbor Laboratory Press, NY.

Neuberger,M.S. and Hartley,B.S. (1979) *J. Mol. Biol.,* **132**, 435-470.

Platt,T. (1981) *Cell,* **24**, 10-23.

Rosenberg,M. and Court,D. (1979) *Annu. Rev. Genet.,* **13**, 319-353.

Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B. (1980) *J. Mol. Biol.,* **143**, 161-178.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA,* **74**, 5463-5467.

Shamanna,D.K. and Sanderson,K.E. (1979a) *J. Bacteriol.*, **139**, 64-70.
Shamanna,D.K. and Sanderson,K.E. (1979b) *J. Bacteriol.*, **139**, 71-79.
Sherman,F. and Stewart,J.W. (1983) *Molecular Biology of Yeast*, published by Cold Spring Harbor Laboratory Press, NY.
Shine,J. and Dalgarno,L. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 1342-1346.
Shine,J. and Dalgarno,L. (1975) *Nature*, **254**, 34-38.
Staden,R. (1977) *Nucleic Acids Res.*, **4**, 4037-4051.
Staden,R. (1978) *Nucleic Acids Res.*, **5**, 1013-1015.
Staden,R. (1979) *Nucleic Acids Res.*, **6**, 2601-2610.
Tinoco,I., Borer,P., Dengler,B., Levine,M., Uhlenbeck,O., Crothers,D. and Gralla,J. (1973) *Nature New Biol.*, **246**, 40-41.
Wang,P.Y., Shopsis,C. and Schneider,H. (1980) *Biochem. Biophys. Res. Commun.*, **94**, 248-254.

Received on 19 December 1983