Purification and Cloning of a Ketoreductase used for the Preparation of Chiral Alcohols

Ronald L. Hanson,* Steven Goldberg, Animesh Goswami, Thomas P. Tully, Ramesh N. Patel

Process Research and Development, Bristol-Myers Squibb, One Squibb Drive, New Brunswick, NJ 08903, USA
Fax: (+1)-732-227-3994, e-mail: ronald.hanson@bms.com

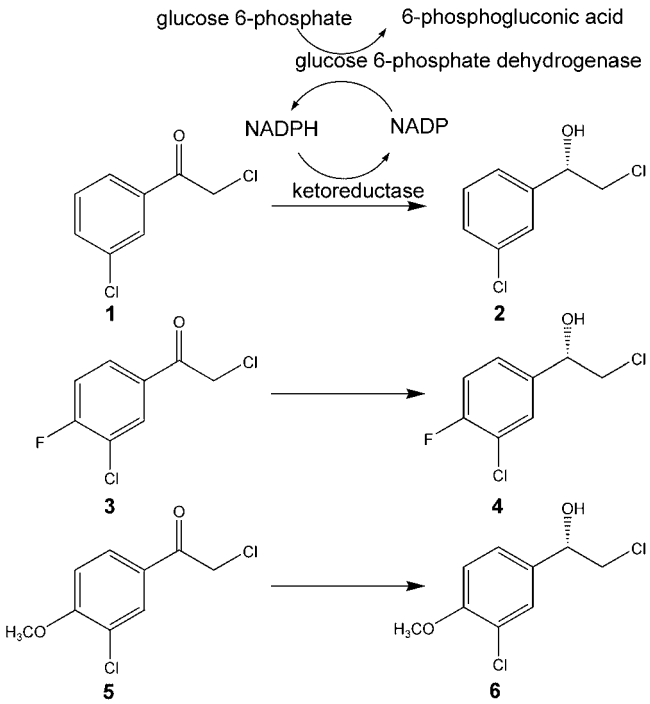
Received: January 24, 2005; Accepted: March 30, 2005

Abstract: The synthesis of the leading candidate compound in an anticancer program required (*S*)-2-chloro-1-(3-chlorophenyl)-ethanol as an intermediate. Other possible candidate compounds used analogues of the *S*-alcohol. Of 119 microbial cultures screened for reduction of the corresponding ketone to the *S*-alcohol, *Hansenula polymorpha* ATCC 58401 (73.8% ee) and *Rhodococcus globerulus* ATCC 21505 (71.8% ee) had the highest enantioselectivity for producing the desired alcohol. A ketoreductase from *Hansenula polymorpha*, after purification to homogeneity, gave the *S*-alcohol with 100% ee. Amino acid sequences from the purified enzyme were used to design PCR primers for cloning the ketoreductase. The cloned ketoreductase required NADP(H), had a subunit molecular weight of 29,220 and a native molecular weight of 88,000. The cloned ketoreductase was expressed in *E. coli* together with a cloned glucose 6-phosphate dehydrogenase from *Saccharomyces cerevisiae* to allow regeneration of the NADPH required by the ketoreductase. An extract of *E. coli* containing the two recombinant enzymes was used to reduce 2-chloro-1-(3-chloro-4-fluorophenyl)-ethanone and two related ketones to the corresponding *S*-alcohols. Intact *E. coli* cells provided with glucose were used to prepare (*S*)-2-chloro-1-(3-chloro-4-fluorophenyl)-ethanol in 89% yield with 100% ee.

Keywords: alcohols; enzyme catalysis; gene expression; *Hansenula polymorpha*; ketoreductase; NADPH regeneration

Introduction

Chiral alcohols are frequently required as synthons for drugs or drug candidate compounds. Preparation by enzymatic resolution has been often described, but unless a dynamic resolution process is developed, this approach is limited to a maximum 50% yield. Reduction of ketones to chiral alcohols by dehydrogenases has also been widely used,[1] and has the advantage of possible 100% yields. Ketoreductases are found in many microbial strains and can give either enantiomer of the alcohol but whole cells may not give high ee because of competing enzymes with lower or opposite enantioselectivity. Usually purification, cloning and overexpression of the desired ketoreductase is required to achieve high ee and productivity. The synthesis of the leading candidate compound in an anticancer program required (*S*)-2-chloro-1-(3-chlorophenyl)-ethanol (**2** in Scheme 1) as an intermediate. Other possible candidate compounds used analogues of the *S*-alcohol. We report here on the identification of strains which reduce the precursor ketone to the desired *S*-alcohol and on the purification, cloning and overexpression of the corresponding ketoreductase.



Scheme 1. Ketones reduced by cloned ketoreductase.

 DOI: 10.1002/adsc.200505045

Cofactor regeneration is a requirement to reduce cost and allow the use of a physiological concentration of the NAD(P)H cofactor. Yeast formate dehydrogenase is commonly used for NADH regeneration, and glucose dehydrogenase from *Bacillus* species may be used for either NADH or NADPH regeneration. Mutated formate dehydrogenases from *Pseudomonas*[2] and *Saccharomyces cerevisiae*[3] that are more active with NADPH have also been described. It is advantageous to clone and express both the ketoreductase and cofactor regeneration enzymes in the same cell. A useful system for production of chiral alcohols has been developed by Kataoka et al.[4] These workers expressed carbonyl reductases together with glucose dehydrogenase from *Bacillus megaterium* in *E. coli* to produce various chiral alcohols. Glucose 6-phosphate dehydrogenase can also be used for NADPH regeneration, and indeed that is one of the major physiological functions of the enzyme. The crystalline enzyme from *Saccharomyces carlsbergensis* was reported to have a specific activity of 676 μmoles NADPH formed per minute per mg protein[5] which is considerably higher than other enzymes used for cofactor regeneration. Synthesis of chiral alcohols using glucose 6-phosphate dehydrogenase for NADPH regeneration has been reported,[6,7] but because of the high cost of glucose 6-phosphate, the use of this method is usually limited to small-scale reactions such as the production of drug metabolites by P450 monooxygenases. Since glucose is transported into *E. coli via* a phosphotransferase system that uses phosphoenol pyruvate as a phosphoryl donor to convert glucose to glucose 6-phosphate,[8] we have tried to exploit this system to use glucose for NADPH regeneration by overexpressing glucose 6-phosphate dehydrogenase from *Saccharomyces cerevisiae* in *E. coli* along with the cloned ketoreductase.

Results and Discussion

Screening

119 strains were screened for conversion of chloroketone **1** to (*S*)- chloroalcohol **2**. Most of the cultures gave predominantly the *R*-alcohol with the highest ee being 97.6%. Of 23 strains that produced the *S*-alcohol *Hansenula polymorpha* ATCC 58401 (73.8% ee) and *Rhodococcus globerulus* ATCC 21505 (71.8% ee) gave the highest ee. Almost all strains screened reduced the ketone to the alcohol, and even excess glucose dehydrogenase reduced the ketone to the alcohol. The coupled enzyme assay and preparative procedures use glucose 6-phosphate dehydrogenase to regenerate the NADPH required for the reduction to avoid the ketone reduction by glucose dehydrogenase (which decreased the ee).

Enzyme Purification and Sequencing

Hansenula polymorpha ATCC 58401 cells (2.088 kg) were obtained from a 50-L fermentation. Using a whole cell assay with 1 mg/mL chloroketone, activity was highest 37 h after the start of the fermentation (72% conversion in 1 h with 80% ee). Activity when cells were harvested after 48 h was 32% conversion in 1 h with 85% ee. Purification of the *Hansenula* ketoreductase was carried out to obtain amino acid sequences necessary for polymerase chain reaction (PCR) amplification of the corresponding gene. During initial purification attempts the activity was not stable, despite the addition of yeast protease inhibitor cocktail, 10% glycerol and 1 mM dithiothreitol. Of various additives tried for stabilization, only NADP and NADPH were effective (NAD was not effective). Therefore 0.5 mM NADP was added to column buffers. The ketoreductase from *Hansenula polymorpha* SC 13824 was purified 112-fold from an extract using Phenylsepharose, Q-Sepharose and Sephacryl S-100. FPLC with a unoQ column was used as a final step to obtain a single band on an SDS gel that was blotted to PVDF for sequencing. A summary of the purification is shown in Table 1. Although the initial extract gave *S*-alcohol with 90% ee, after the Q-Sepharose chromatography step the enzyme preparation produced *S*-alcohol with 100% ee. The purified protein required NADP and was not active with NAD. The subunit molecular weight determined with an Agilent Bioanalyzer 2100 Protein 200 Labchip was 27 kd.

An N-terminal sequence (H2N-DKLPTEAPQLP) was obtained which showed sequence homology with sorbitol dehydrogenase from *Candida albicans*. Estimated protein concentration from the Agilent Bioanalyzer 2100 indicated that 200 pmoles of protein per lane were provided for sequencing. However, the N-terminal sequence contained only about 2.5 pmoles, indicating most of the N-terminal amino acid may be blocked. A tryptic digest gave several peaks which

Table 1. Purification of the ketoreductase from extract of *Hansenula polymorpha* ATCC 58401.

Step	Volume [mL]	Activity [units]	Protein [mg]	Specific activity [U/mg]
Extract	380	72.2	629.3	0.114
Phenyl Sepharose	48	31.3	22.8	1.37
Q Sepharose	24	15.0	2.04	7.35
Sephacryl S-100	24	6.6	0.516	12.79

contained an estimated 50–100 pmoles of peptide and sequences of two of the internal peptides (AIMNTNLDGVYYCAK and SLAMEWVGFAR) were obtained. These sequences also showed homology with dehydrogenases in the data bases.

Cloning

Degenerate primers for the N-terminal and internal sequences were used in PCR reactions to clone 400 and 600 base pair (bp) fragments of the ketoreductase gene from *Hansenula polymorpha* DNA. The 400 bp fragment was used to prepare a digoxigenin-labeled probe which bound to a 4 kilobase (kb) band from a *Xho*I digest of *Hansenula polymorpha* ATCC 58401 chromosomal DNA. A library was prepared in *E. coli* containing DNA from the 4 kb region of the *Xho*I digest ligated to the pZErO-2 vector, and the probe was used to identify *E. coli* colonies carrying the homologous sequence. Sequencing of the plasmid prepared from one of the hybridizing colonies showed that the insert contained an open reading frame homologous to known ketoreductases and that the sequences corresponding to the amino terminal and two internal peptides of the purified enzyme were present (Figure 1). Primers that included *Nde*I and *Sma*I sites before the start and stop sites of the insert, respectively, were used to amplify the ketoreductase gene. The amplified fragment was ligated into pBMS2000 for expression in *E. coli* BL21Star(DE3) under the control of the *tac* promoter. After growth of the cells and induction with 1 mM isopropyl β-D-thiogalactoside (IPTG), a sonicated extract was prepared that contained a predominant 27 kd protein (Figure 2, lanes 4 and 9). In an activity assay, the extract gave complete conversion of chloroketone **1** to the chloro-*S*-alcohol **2** with 100% ee. An extract of *E. coli* BL21Star(DE3) cells lacking the plasmid produced only a trace of the chloroalcohol HPLC peak.

```
      M  N  I  I  G  N  Y  D  K  L  P  T  E  A  P  Q
  1   ATGAACATTATCGGAAATTACGACAAGCTACCAACCGAGGCTCCTCAA
      L  P  S  N  V  F  S  L  F  S  L  K  G  K  V  A
 49   TTGCCTTCCAACGTTTTCAGCCTGTTCTCCCTGAAAGGCAAGGTGGCC
      S  I  T  G  G  S  T  G  I  G  L  A  V  A  E  A
 97   AGCATTACTGGTGGCTCGACAGGAATTGGTCTGGCTGTGGCAGAAGCG
      Y  A  Q  A  G  A  D  V  A  I  W  Y  N  S  T  N
145   TATGCTCAGGCAGGCGCAGACGTGGCCATCTGGTACAACAGCACAAAC
      A  D  H  E  A  E  R  L  S  K  T  Y  G  I  R  A
193   GCTGACCACGAAGCTGAGAGGCTGTCCAAGACGTACGGGATCCGTGCC
      K  A  Y  K  C  A  V  G  D  F  D  Q  V  K  A  T
241   AAGGCTTACAAGTGCGCAGTGGGCGACTTTGACCAGGTCAAGGCCACG
      I  D  A  I  E  S  D  F  G  T  I  H  I  F  V  A
289   ATCGATGCCATTGAGTCTGACTTTGGCACGATTCACATTTTTGTTGCA
      N  A  G  I  G  S  Q  S  V  P  V  I  D  A  S  L
337   AATGCGGGGATTGGCTCCCAATCGGTGCCTGTGATCGATGCGTCGCTG
      E  K  Y  R  A  I  M  N  T  N  L  D  G  V  Y  Y
385   GAAAAATACCGGGCAATCATGAACACGAATTTGGACGGCGTGTACTAC
      C  A  K  C  V  G  P  I  F  K  K  H  G  K  G  S
433   TGCGCCAAGTGCGTGGGTCCAATTTTCAAGAAGCACGGCAAGGGTTCC
      F  I  I  T  T  S  Q  A  A  H  I  V  T  A  H  V
481   TTTATCATCACCACCTCACAGGCAGCCCATATTGTCACGGCTCACGTG
      W  Q  A  A  Y  N  A  S  K  A  A  C  I  Q  I  A
529   TGGCAAGCGGCTTACAACGCCAGCAAGGCAGCGTGCATCCAGATTGCC
      K  S  L  A  M  E  W  V  G  F  A  R  V  N  T  I
577   AAGAGTCTGGCAATGGAATGGGTCGGCTTCGCCCGTGTCAATACGATC
      S  P  G  Y  I  V  T  P  I  S  K  D  V  P  N  E
625   TCTCCAGGGTACATTGTCACCCCTATCTCGAAAGATGTGCCTAACGAG
      E  K  V  K  W  C  T  L  I  P  M  G  R  E  G  L
673   GAGAAAGTCAAGTGGTGCACGTTGATCCCAATGGGCAGAGAGGGGCTT
      P  Q  E  L  V  G  A  Y  L  Y  F  A  S  D  A  S
721   CCTCAAGAGCTTGTGGGCGCATACCTGTACTTTGCGTCGGACGCCTCA
      T  F  T  T  G  A  D  L  I  I  D  G  G  Y  C  C
769   ACCTTCACCACCGGCGCTGACCTCATCATTGACGGTGGTTATTGCTGC
      P  *
817   CCATAA
```

Figure 1. Sequence of ketoreductase gene. Peptides from amino acid sequencing are underlined in bold.

Molecular Weight of Native Ketoreductase

The molecular weight of the expressed ketoreductase in the extract was estimated as 88,094 by Superdex-200 gel filtration chromatography. The peak fraction from the

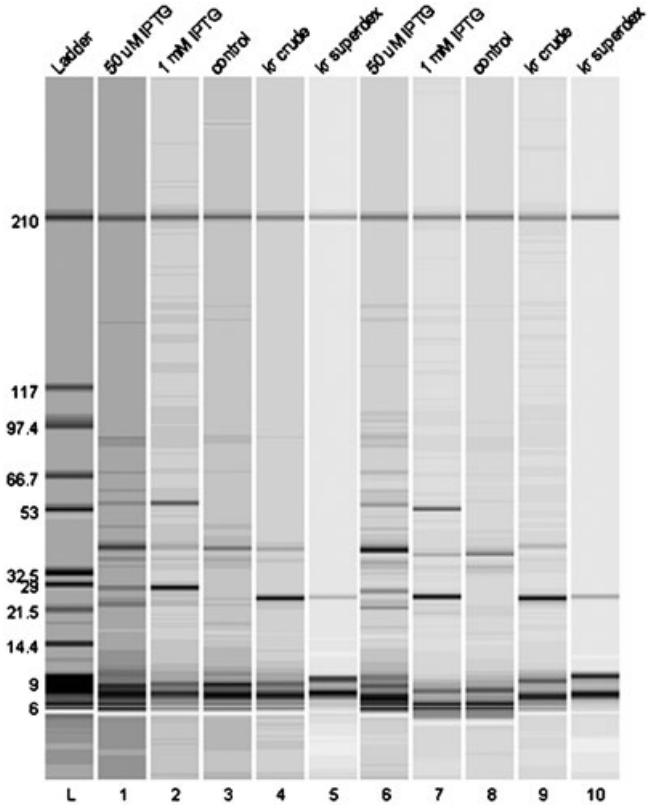


Figure 2. Analysis of expression of cloned enzymes in *E. coli* extracts. Lanes 1, 6, 2, 7: extract of strain with ketoreductase (27 kd) and glucose 6-phosphate dehydrogenase (55 kd). Lanes 1, 6 were induced with 50 μM IPTG. Lanes 2, 7 were induced with 1 mM IPTG. Lanes 4, 9: extract of strain with ketoreductase only induced with 1 mM IPTG. Lanes 5, 10: ketoreductase peak from Superdex-200 chromatography. Lanes 3, 8: extract of control *E. coli* (no plasmid). Extracts were from 10% w/v sonicates and were analyzed with an Agilent 2100 Bioanalyzer Protein 200 Labchip.

column contained a single 27 kd subunit when analyzed with a Protein 200 Labchip (Figure 2, lanes 5 and 10). From the amino acid composition derived from the gene sequence, the subunit molecular weight is 29,220, suggesting that the native enzyme is a trimer.

Expression of *Hansenula polymorpha* Ketoreductase and *Saccharomyces cerevisiae* Glucose 6-Phosphate Dehydrogenase in *E. coli*

The cloned ketoreductase was copied from the plasmid by PCR. *Sma*I and *Hind*III sites were added upstream of the promoter and a *Sma*I site was added at the 3′ end *via* the primers. The ketoreductase gene was ligated at a *Sma*I site into pBMS2000-SCGD which contains a cloned *S. cerevisiae* glucose 6-phosphate dehydrogenase gene. Restriction digests with *BamH*I and *Nde*I were used to identify plasmids with the two genes in the same orientation, which was thought to be a desired situation for a higher transcription rate. A restriction site map of the construct is shown in Figure 3. One of the plasmids with the desired orientation was used to transform BL21Star(DE3) *E. coli* for expression. This strain was designated *E. coli* SC16529. The cells were induced with either 50 μM or 1 mM IPTG. Expression of both enzymes was 2- to 3-fold better at 1 mM IPTG (Figure 2, lanes 1, 6 *versus* lanes 2, 7). Sonicated extracts of the cells with either concentration of inducer completely converted chloroketone **1** to chloroalcohol **2** in buffer containing NADP and glucose 6-phosphate.

Alternative Substrates

Two other ketones (**3** and **5** in Scheme 1, intermediates for potential drug candidates) were reduced equally as fast as **1** by the 1 mM IPTG-induced *E. coli* extract and gave single peaks with the chiral HPLC method used for **2**. The alcohol **4** produced by the enzymatic reduction was also shown to be the *S*-enantiomer by comparison with an authentic standard. No chiral marker was available for the product alcohol **6**.

Preparative Batches of Alcohol 4

E. coli SC16529 was grown in a fermentor and the expression of the enzymes was induced with 1 mM IPTG as described in the Experimental Section to provide an enzyme source for preparative batches. Using an extract of the recombinant *E. coli*, batches of 1, 2, and 20 g ketone **3** were reduced. The product alcohol had 100% ee in each batch. Conditions for the 20 g batch were: 20 g/L ketone **3** (97 mM), 40 g/L glucose 6-phosphate (142 mM), 191 mg/L NADP (0.25 mM), 0.1 M potassium phosphate buffer, pH 7, 5% methanol and 5% w/v *E. coli* as a microfluidized and centrifuged extract. The yield of **4** (isolated, but not purified) was 91% with 100% ee. The amount of extract used contained 181,000 units of ketoreductase and 47,000 units of glucose 6-phosphate dehydrogenase and was calculated to be capable of reducing the ketone in 2.05 minutes. The ketone was 96% depleted in 30 minutes, and when the reaction was stopped after 5.5 h, 0.5% ketone remained. Enzyme assays of samples from a 2-g batch indicated that both enzymes were inactive 45 minutes after the start of the batch, probably from the addition of methanol.

A procedure using intact cells of recombinant *E. coli* with glucose supplied as a precursor of glucose 6-phosphate was developed. The final best conditions used for batches of 2- and 15-g were: 20 mg/mL ketone **3** (97 mM), 40 mg/mL glucose (222 mM), 0.153 mg/mL NADP (0.2 mM), 0.1 M potassium phosphate buffer, pH 7, and 10% w/v *E. coli*. The isolated yield of unpurified **4** was 89% with 100% ee. When the ketone was added in methanol (5% final concentration) or when NADP was omitted, reaction was slow and incomplete.

Initially, the product alcohol was isolated using ethyl acetate extraction, then MTBE extraction was used to get less emulsion during extraction. Solid phase extraction using the resin XAD-16 as described in the Experimental Section was a more suitable procedure and gave good quality product in high yield.

Conclusion

A ketoreductase from *Hansenula polymorpha* was purified, cloned and overexpressed in *E. coli*, and was used to prepare (*S*)-2-chloro-1-(3-chloro-4-fluorophenyl)-ethanol in high yield and ee. Glucose 6-phosphate dehydrogenase from *Saccharomyces cereviseae* was also expressed in the same strain to allow glucose to be used in whole cells for NADPH regeneration. Further metabolic engineering of the strain could be used to increase the production of phosphoenol pyruvate needed for glucose phosphorylation and to enhance the remaining enzymes needed for the phosphogluconate oxidative pathway.

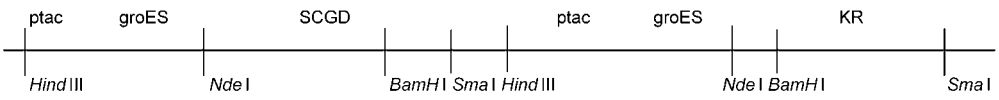


Figure 3. Ketoreductase (KR)/glucose-6-phosphate dehydrogenase (SCGD) construct.

Experimental Section

Screening

Z49 medium contained 3% trypticase soy broth powder (BBL brand). F7 medium contained 1% malt extract, 1% yeast extract, 0.1% peptone and 2% dextrose adjusted to pH 7. Cultures were grown in 2 mL medium (16 strains on Z49, 103 strains on F7) in multiwell plates shaken at room temperature. After 24 h, 20 μL of a 100 mg/mL solution of chloroketone **1** in methanol were added to each well and the incubation was continued for 3 days. Samples were extracted with 2 mL ethyl acetate per mL broth. The extract (1 mL) was dried at 40 °C under a nitrogen stream and the residue was dissolved in 1 mL 2-propanol for HPLC analysis of ketone and alcohol concentration and ee of *S*-alcohol.

Growth of *Hansenula polymorpha* ATCC 58401 and Preparation of Cell-Free Extract

Note: Strains deposited as *Hansenula polymorpha* are now called *Pichia angusta* by the American Type Culture Collection. Cells were grown in 50 liters F7 medium for 48 h starting from a 0.5 liter inoculum. Airflow was 1 vvm, temperature was 28 °C, agitation was 500 rpm and SAG antifoam was used to control foaming. Cell paste (2.088 kg) was recovered by centrifugation and stored at −70 °C. Cells (20% w/v) were suspended in 50 mM potassium phosphate buffer pH 7, containing 10% glycerol, 1 mM dithiothreitol, 1% yeast protease inhibitor cocktail (Sigma) and disrupted by 10 passages through a microfluidizer. Unbroken cells and debris were removed by centrifugation of the extract at 20434 × g for 20 min. The extract was stored at −20 °C until used for purification.

Purification of Ketoreductase

Extract (380 mL) was applied to 50 mL Amersham/Pharmacia Phenyl Sepharose 6 fast flow in a 26 mm diameter column at 3 mL/min. The column was packed in 20 mm tris (chloride) pH 7.4 containing 10% glycerol, 1 mM dithiothreitol and 0.5 mM NADP (buffer A). The column was washed at 1 mL/min with 120 mL buffer A, then 300 mL buffer A containing 1% triton X-100 (Pierce Surfact-Amps® X-100). The 48 mL peak of activity was applied to 30 mL Amersham/Pharmacia Q Sepharose fast flow in a 16-mm diameter column packed in buffer A and eluted with a 240-mL 0 to 0.2 M NaCl gradient in buffer A at a flow rate of 2 mL/min. The activity peak fractions containing 24 mL were concentrated to about 6 mL using an Amicon YM10 membrane in a stirred cell. The concentrated enzyme was passed at a flow rate of 1 mL/min through 450 mL Amersham/Pharmacia Sephacryl S-100 packed in a 2.6 × 100 cm column with 50 mM potassium phosphate buffer pH 7 containing 10% glycerol, 1 mM dithiothreitol, 0.5 mM NADP and 0.15 M NaCl. The peak two fractions containing 24 mL were concentrated and diafiltered into buffer A with an Amicon Centricon-10, loaded onto a Biorad Uno Q1 column (1.3 mL), then eluted with a 16-mL 0 to 0.25 M NaCl gradient in buffer A at a flow rate of 1 mL/min. The activity peak fractions containing 1 mL were concentrated with an Amicon Microcon-10 to 0.1 mL then subjected to SDS gel electrophoresis and blotted onto a PVDF membrane. Sequencing of the N-terminus and two internal peptides was done by Argo Bioanalytica. Purity and molecular weight of the enzyme subunit were determined with an Agilent Bioanalyzer 2100 using a Protein 200 Labchip.

Enzyme Assays

Prior to cloning the ketoreductase, enzyme assays were performed using HPLC. To evaluate activities of whole cells during fermentations, cells were pelleted from 0.5 mL broth with a microfuge and resuspended in 0.5 mL 50 mM potassium phosphate buffer pH 7 containing 10 mg/mL glucose and 0.5 mg **1** dissolved in 5 μL methanol. For cell extracts and column fractions, ketoreductase assays contained in 0.5 mL: 40 mM potassium phosphate buffer pH 7, 0.8 mM NADP, 8.5 mM sodium glucose 6-phosphate, 1.6 units glucose 6-phosphate dehydrogenase from baker's yeast, 0.5 mg **1** dissolved in 5 μL methanol and ketoreductase sample. Assays were carried out in microfuge tubes at room temperature on a 12 rpm end-over-end rotator. Reactions were terminated (typically after 30 min) by addition of 0.5 mL acetonitrile for quantitation by reverse phase HPLC or by extraction with 1 mL ethyl acetate for ee measurements by chiral HPLC. Better results were obtained for chiral chromatography when the samples were dissolved in mobile phase instead of 2-propanol. Activities were calculated after HPLC analysis as 1 μmole/min = 1 unit.

After cloning, spectrophotometric enzyme assays were used to measure activity in extracts of recombinant *E. coli*. Ketoreductase assays contained in 1 mL: 0.1 M potassium phosphate buffer pH 7, 0.3 mM NADPH, and 0.1 mg ketone **3** in 0.05 mL methanol. The reaction was started with cell extract (diluted into 50 mM potassium phosphate buffer, pH 7, containing 1 mM dithiothreitol). Absorbance decrease at 340 nm was used to calculate the enzyme activity as 1 μmole/min = 1 unit.

Glucose 6-phosphate dehydrogenase assays contained in 1 mL: 0.1 M potassium phosphate buffer, pH 7, 0.3 mM NADP and 0.5 mM sodium glucose 6-phosphate. The reaction was started with cell extract (diluted into 50 mM potassium phosphate buffer, pH 7, containing 1 mM dithiothreitol). Absorbance increase at 340 nm was used to calculate the enzyme activity as 1 μmole/min = 1 unit.

HPLC Analysis

Quantitation was performed with a YMC Pak ODS A 3 μ 15 × 0.6 cm column. The mobile phase was a gradient of 20% acetonitrile/80% water to 90% acetonitrile/10% water from 0 to 12 min, flow rate was 1 mL/min, column temperature was 40 °C, detection was with a DAD set at 220 nm and injection volume was 5 μL. Retention times were: chloroalcohol **2** 9.9 min, chloroketone **1** 11.1 min.

Chiral separations were performed with a Chiralpak AD-H 25 × 0.46 cm column (Daicel Chemical Industries, Ltd.). The mobile phase was 95% hexane/5% ethanol, flow rate was 1 mL/min, column temperature was 18 °C, detection was with a DAD set at 220 nm and injection volume was 10 μL. Retention times were: *S*-enantiomer of **2** 10.4 min, *R*-enantiomer of **2** 13.1 min, chloroketone 11.0 min. Authentic chiral markers

were prepared by Bristol-Myers Squibb chemists by Corey reduction of the ketones.

Cloning of the Ketoreductase Gene

Degenerate primers were prepared based on the N-terminal (sense) and two internal amino acid sequences (antisense) of the ketoreductase. They were used in PCR with chromosomal DNA from *Hansenula polymorpha* SC 13824 and gave 400-bp and 600-bp products. The 400-bp product was ligated to PCR4-TOPO as recommended by the manufacturer (Invitrogen) and transformed into TOP10 *E. coli* (Invitrogen) by electroporation. The sequence of the insert was obtained from the Bristol-Myers Squibb sequencing facility and showed homology with short chain dehydrogenases/reductases. The 400-bp PCR product was used as template to prepare a digoxigenin-labeled probe (Roche Biochemicals) to screen five restriction enzyme digests of chromosomal DNA from *Hansenula polymorpha* ATCC 58401. All five digests contained a band that bound the probe. The probe hybridized to a *Xho*I fragment of about 4 kbp which was thought large enough to contain the full-length ketoreductase gene. DNA from this region was purified and inserted into the pZErO-2 vector (Invitrogen), then used to transform DH10B (Invitrogen) cells by electroporation. Two colonies that contained the insert were identified by hybridization with the probe and confirmed to contain the 400 bp sequence by colony PCR. A pZErO-2 plasmid that contained the insert was prepared, and random primers were inserted with the GPS-1 Tn7 transposon-based procedure (New England Biolabs) for use in sequencing the plasmid. Sequencing showed that the insert contained the full length gene for the ketoreductase, and the sequences corresponding to the amino terminal and two internal peptides obtained from the purified enzyme were present. Using PCR with the appropriate primers and the plasmid containing the full length gene as template, an *Nde*I site was introduced at the start of the gene and an *Sma*I site was introduced after the stop codon. The ends of the PCR product were digested with these two restriction enzymes then ligated into pBMS2000 cut with the same enzymes. The ligation product was used to transform DH5α-T1 (Invitrogen) chemically-competent *E. coli* cells, and colonies containing the ketoreductase gene insert were identified by PCR. Plasmids were isolated from two of the colonies and one of the plasmids was used to transform BL21Star(DE3) *E. coli* cells (Invitrogen) for expression. One of the colonies was grown at 30 °C in MT5 medium (composition given below), and induced at OD600 = 0.99 with 1 mM IPTG. After 4 hours induction, cells were harvested and stored frozen. Extracts (10% w/v) were prepared by sonication and found to contain a predominant 27 kd protein with an Agilent 2100 Bioanalyzer Protein 200 Labchip. In an activity assay, the extract gave complete conversion of chloroketone **1** to the *S*-alcohol **2** with 100% ee.

Estimation of the Molecular Weight of the Ketoreductase Expressed in *E. coli*

A sample of the extract (0.1 mL) prepared above in 50 mM potassium phosphate buffer pH 7 containing 10% glycerol and 1 mM dithiothreitol was applied to a Superdex 200 column (10 × 300 mm) and run at 0.5 mL/min using an AKTA FPLC. The column buffer was 50 mM potassium phosphate pH 7 containing 10% glycerol, 0.1 M NaCl and 1 mM dithiothreitol. Protein molecular weight standards were aldolase (158000), bovine serum albumin (67000), ovalbumin (43000), chymotrypsinogen A (25000) and ribonuclease (13700).

Construction of Strain with Ketoreductase and Glucose 6-Phosphate Dehydrogenase

The cloned ketoreductase gene was transferred to pBMS2000-SCGD which contains a glucose 6-phosphate dehydrogenase gene for NADPH regeneration. The glucose 6-phosphate dehydrogenase gene was PCR cloned from *Saccharomyces cerevisiae* DNA using primers based on a published sequence[9] as an *Nde*I-*Eco*RI fragment and expressed in pBMS2000 by Thomas Franceschini and Li You (Bristol-Myers Squibb). Attempts to cut the ketoreductase gene from plasmid pBMS2000 with *Hind*III and *Sma*I were not successful, apparently because the *Sma*I site was lost upon cloning. Therefore, the gene and promoter were copied from the plasmid by PCR. *Sma*I and *Hind*III sites were added upstream of the promoter and an *Sma*I site was incorporated into the primer homologous to the 3′ end of the ketoreductase gene. The amplified product was trimmed with *Sma*I then ligated into pBMS2000-SCGD (plasmid containing cloned glucose 6-phosphate dehydrogenase from *Saccharomyces cerevisiae*) cut with *Sma*I. The ligation product was used to transform DH5α-T1 *E. coli* cells. Plasmids were isolated from seven of the transformants and orientations of the two cloned genes were checked by cutting with *Hind*III, *Bam*H1, or *Nde*I. One of the *Hind*III sites was apparently lost and no conclusion could be drawn from that digestion. However the *Bam*H1 and *Nde*I digestions were consistent with three of the plasmids having the two genes in the same orientation (desired situation for higher transcription rates) and one of the plasmids having the genes with opposite orientation (undesired). One of the plasmids with the desired orientation was used to transform BL21Star(DE3) *E. coli* for expression. This strain was designated *E. coli* SC16529.

E. coli SC16529 (pBMS2000-SCGD-KR) Fermentation in 15-Liter Tanks

MT5 inoculum medium contained: 2.0% Yeastamin; 4.0% glycerol; 0.6% sodium phosphate, dibasic; 0.3% potassium phosphate, monobasic; 0.125% ammonium sulfate; 0.0246% magnesium sulfate heptahydrate (added post-autoclaving from a filter-sterilized solution); 0.005% kanamycin (added post-autoclaving from a filter-sterilized solution); pH was adjusted to 7.2 prior to autoclaving at 121 °C.

MT5-M1 fermentor medium contained: 2.0% Quest Hy-Pea; 1.85% Tastone-154; 4.0% glycerol; 0.6% sodium phosphate, dibasic; 0.125% ammonium sulfate; 0.04% UCON antifoam; 0.0246% magnesium sulfate heptahydrate (added post-sterilization from a filter-sterilized solution). Medium in the fermentor was batched with deionized water to yield a final volume of 15 liters after post-sterilization additions and feeds. The pH was adjusted to 7.4 prior to sterilization.

A frozen vial of *E. coli* SC16529 was thawed and transferred to a 4-liter flask containing 1 liter of MT5 medium. The flask

was incubated at 30 °C and 250 rpm for *ca.* 24 h, at which time the OD_{600} of the F1 stage growth was 6.4 units/cm. The fermentor, containing 15 liters of MT5-M1 medium, was then inoculated with sufficient broth (*ca.* 820 mL) to yield an optical density in the tank of 0.35 units/cm. Induction was performed approximately 2.5 h post-inoculation. IPTG, which had been dissolved in 500 mL water and filter-sterilized through a 0.2 um filter, was added aseptically to the tank to yield a final concentration of 1.0 mM. The tank was run at 30 °C with an agitator speed of 500 rpm and airflow of 1 vvm. Ammonium hydroxide was used to maintain pH at 7.0 throughout the run. The tank was harvested 21 h after inoculation. At this time, the broth was subcooled and centrifuged, followed by a wash with 10 mM pH 7.0 sodium phosphate buffer and a second centrifugation. The final yield was 296 g cell paste which was frozen and stored at −70 °C until needed for preparative biotransformation.

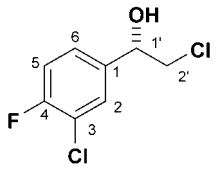
Preparation of 4 using Cell Extract

E. coli SC16529 (80 g wet cell weight) was suspended to 400 mL in 50 mM potassium phosphate buffer pH 7 containing 1 mM dithiothreitol, passed twice through a microfluidizer at 12000 psi and centrifuged for 15 min at 23,424 × g. The supernatant was used as a source of ketoreductase and glucose 6-phosphate dehydrogenase. To a 2-L jacketed reactor magnetically stirred and maintained at 30 °C were added: sodium glucose 6-phosphate (40 g, 142 mmoles), 650 mL water and 100 mL 1 M potassium phosphate buffer pH 7. The solution was adjusted to pH 7 by addition of about 12.5 mL of 10 N NaOH. Cell extract (200 mL containing 724 U/mL ketoreductase and 188 U/mL glucose 6-phosphate dehydrogenase) and NADP (153 mg, 0.2 mmoles) were added to the solution. Ketone **3** (20 g, 97 mmoles) dissolved in 50 mL methanol was slowly pipetted into the stirred solution. The progress of the reduction was followed by HPLC. After 4.5 h the pH was adjusted from 6.35 to 7 by addition of about 7 mL 10 N NaOH and an additional 50 mL cell extract and 38 mg NADP (0.05 mmoles) were added.

After 5.5 h total reaction time, 200 g of washed XAD-16 resin were added and the mixture was stirred with a mechanical stirrer overnight (16 h). The adsorption was followed by HPLC. After 16 hours only about 1.6% of alcohol **4** remained in the aqueous layer. The mixture was filtered through a 40 mesh stainless steel screen to collect the XAD-16. The XAD-16 was washed with about 3000 mL water, then ethyl acetate (500 mL) was added, the mixture was stirred with a mechanical stirrer for 60 minutes, then filtered through the stainless steel screen mesh. The filtrate contained about 150 mL water and ethyl acetate (500 mL). The two layers were separated. Only 1% of the total alcohol was in the water, which was discarded. The desorption of XAD-16 was nearly complete with two more 500-mL portions of ethyl acetate. There was no water layer after the first desorption. The combined ethyl acetate solution was dried over anhydrous Na_2SO_4 and filtered (glass fritted disc, medium porosity). Removal of solvent in the rotary evaporator provided 20 g of an oil. After drying for 24 h in vacuum, 19.6 g of a light yellowish oil were obtained. HPLC showed that the alcohol area percent (AP) was 94.01 and ketone AP was 1.26. The yield of **4** corrected for AP was 91%, and ee was 100%.

NMR of Alcohol

NMR spectra were recorded in $CDCl_3$ solution with TMS as internal standard using a Jeol NMR spectrophotometer (400 MHz for ^1H and 100 MHz for ^{13}C) and δ values are reported. ^1H NMR: δ = 2.86 (1 H, br s, OH), 3.68 (2 H, m, H-2′), 4.84 (1 H, m, H-1′), 7.13 (1 H, m, H-5), 7.22 (1 H, m, H-6), 7.44 (1 H, m, H-2); ^{13}C NMR: δ = 50.99 (C-2′), 73.13 (C-1′), 116.57 (C-6), 116.79 (C-3), 125.78 (C-5), 128.29 (C-2), 136.74 (C-1), 156.27 and 158.74 (C-4).



Preparation of 4 using Whole Cells

E. coli SC16529 (75 g wet cell weight) was dispersed in a solution containing glucose (30 g, 167 mmoles), 675 mL water and 75 mL 1 M potassium phosphate buffer pH 7 by using an Ultraturrax homogenizer. Dry cell weight was 34.6% of the wet cell weight of the *E. coli* cell paste. The cell suspension was added to a 1-L jacketed reactor, mechanically stirred and maintained at 30 °C. NADP (115 mg, 0.15 mmoles) and ketone **3** (15 g, 72.4 mmoles, finely ground with a mortar and pestle) were added to start the reaction. pH 7 was maintained by periodic additions of 10 N NaOH during the first two hours, and later by addition of 1 N NaOH with a pH stat.

After 22 h reaction time, 150 g of washed XAD-16 were added and the mixture was stirred with a mechanical stirrer. After 21 h, HPLC showed about 3% of the alcohol **4** remaining in the aqueous layer. After 22 h, the mixture was filtered through a 40 mesh stainless steel screen. The XAD-16 was washed with 2000 mL water, then the product alcohol **4** was eluted with ethyl acetate (300 mL, 3 times). The 50 mL water layer from the first desorption was also back-extracted with 100 mL ethyl acetate. The combined ethyl acetate eluates were dried over anhydrous Na_2SO_4, filtered, and solvent was removed with a rotary evaporator. After drying for 20 h in vacuum, 14.8 g of a light yellowish oil were obtained. HPLC showed that the alcohol AP was 91.59 and ketone AP was 2.32. The yield of **4** corrected for AP was 89% and ee was 100%.

References

[1] M.-R. Kula, in: *Stereoselective Biocatalysis*, (Ed.: R. N. Patel), Marcel Dekker, New York, **2000**, pp. 839–866.
[2] V. I. Tishkov, A. G. Galkin, V. V. Fedorchuk, P. A. Savitsky, A. M. Rojkova, H. Gieren, M.-R. Kula, *Biotechnol. Bioeng.* **1999**, *64*, 187–193.
[3] A. E. Serov, A. S. Popova, V. V. Fedorchuk, V. I. Tishkov, *Biochem. J.* **2002**, *367*, 841–847.
[4] M. Kataoka, K. Kita, M. Wada, Y. Yasohara, J. Hasegawa, S. Shimizizu, *Appl. Microbiol. Biotechnol.* **2003**, *62*, 437–445.

[5] E. A. Noltmann, C. J. Gubler, S. A. Kuby, *J. Biol. Chem.* **1961**, *236*, 1225–1230.
[6] C.-H. Wong, G. M. Whitesides, *J. Am. Chem. Soc.* **1981**, *103*, 4890–4899.
[7] W. Hummel, *Appl. Microbiol. Biotechnol.* **1990**, *34*, 15–19.
[8] N. D. Meadow, D. K. Fox, S. Roseman, *Annu. Rev. Biochem.* **1990**, *59*, 497–542.
[9] D. Thomas, H. Cherest, Y. Surdin-Kerjan, *EMBO J.* **1991**, *10*, 547–553.

