Reassembled Biosynthetic Pathway for Large-Scale Carbohydrate Synthesis: α-Gal Epitope Producing "Superbug"

Xi Chen, Ziye Liu, Jianbo Zhang, Wei Zhang, Przemyslaw Kowal, and Peng George Wang*[a]

A metabolic pathway engineered Escherichia coli strain (superbug) containing one plasmid harboring an artificial gene cluster encoding all the five enzymes in the biosynthetic pathway of Galα1,3Lac through galactose metabolism has been developed. The plasmid contains a λ promoter, a cl857 repressor gene, an ampicillin resistance gene, and a T7 terminator. Each gene was preceded by a Shine–Dalgarno sequence for ribosome binding. In a reaction catalyzed by the recombinant E. coli strain, Galα1,3Lac trisaccharide accumulated at concentrations of 14.2 mM (7.2 g L^{-1})

in a reaction mixture containing galactose, glucose, lactose, and a catalytic amount of uridine 5′-diphosphoglucose. This work demonstrates that large-scale synthesis of complex oligosaccharides can be achieved economically and efficiently through a single, biosynthetic pathway engineered microorganism.

KEYWORDS:

biosynthesis · gene expression · glycosylation · metabolic engineering · oligosaccharides

Introduction

Recently developed genomic databases have granted us a wealth of information on the biosynthetic pathways of a great number of natural compounds. Although remarkable progress has been made in natural product synthesis by manipulating biosynthetic pathways,[1–5] a variety of valuble compounds still cannot be mass-produced practically. Therefore, a technological link between biosynthetic-pathway databases and mass production would have enormous implications. This is also the case for the production of oligosaccharides and glycoconjugates.

It is well established that the formation of a glycosidic bond in a biological system involves activation of a monosaccharide by attaching it to a nucleotide, followed by the transfer of the monosaccharide from this activated, high-energy sugar–nucleotide donor to an acceptor to form an elongated sugar and a nucleotide by-product (Scheme 1). The sugar–nucleotide conjugate does not appear in the overall, balanced equation (Scheme 1, bottom) and only serves as an intermediate in the glycosylation. Such a biosynthetic pathway was well utilized in the pioneering work by Wong, Haynie, and Whitesides on the in vitro enzymatic synthesis of N-acetyllactosamine with in situ regeneration of uridine 5′-diphosphogalactose (UDP-Gal).[6] Since then, many glycosylation cycles with regeneration of sugar–nucleotides have been developed with either native or recombinant enzymes.[7–10] Indeed, some of these synthetic cycles have been

Sugar-nucleotide biosynthesis pathway:

Overall:

$$\text{HO}-S_1\text{-OH} \ + \ \text{HO}-S_2\text{-OH} \ \longrightarrow \ \text{HO}-S_1\text{-O-}S_2\text{-OH} \ + \ H_2O \quad (1)$$
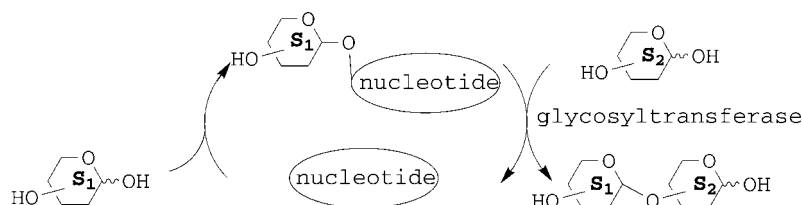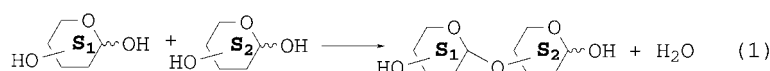
Scheme 1. A general biosynthetic route for the formation of a glycosidic bond.

implemented by industry to produce oligosaccharides on a gram or even kilogram scale.[11, 12]

Recently, a synthesis of oligosaccharides by using a whole-cell approach has emerged. Both Escherichia coli[13, 14] and yeast cells[15] were used. Oligosaccharides were produced in gram scale by an E. coli strain engineered to overexpress a glycosyl-
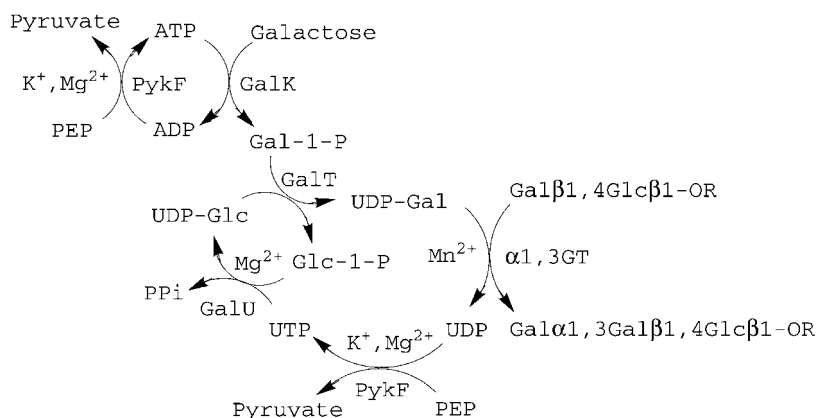
[a] Prof. Dr. P. G. Wang, Dr. X. Chen, Dr. Z. Liu, Dr. J. Zhang, Dr. W. Zhang, P. Kowal
    Department of Chemistry
    Wayne State University
    Detroit, MI 48202 (USA)
    Fax: (+1) 313-577-2554
    E-mail: pwang@chem.wayne.edu

transferase.[16] Elongated oligosaccharides were also obtained by introducing several glycosyltransferases into *E. coli*, whose expressions were controlled by compatible promoters in separated plasmids.[17–19] A more efficient whole-cell approach was presented in the multi-plasmid/strain system developed by Kyowa Hakko Kogyo Co. Ltd. in Japan. The key component in Kyowa Hakko's technology was a *Corynebacterium ammoniagenes* strain, which was engineered to efficiently convert orotic acid into UTP. When combined with *E. coli* strains engineered to overexpress enzymes for sugar–nucleotide biosynthesis, the sugar–nucleotide conjugate accumulated in the reaction solution. By the addition of a recombinant *E. coli* strain overexpressing a glycosyltransferase, a high concentration of oligosaccharide was achieved.[20–25]
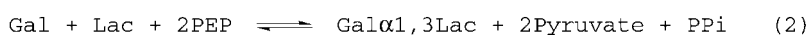
In the present work, we have examined the feasibility of transferring the complex in vitro biosynthetic cycles into a single, product-producing *E. coli* strain containing a plasmid with all the necessary genes for sugar–nucleotide regeneration and oligosaccharide accumulation.

As an example, the methodology was applied to the synthesis of Galα1,3Lac trisaccharide (also called α-Gal epitope). Galα1,3-Gal-terminated oligosaccharide sequences exist on cell-surface glycolipids or glycoproteins in mammals other than humans, apes, and old world monkeys.[26] It is the major antigen responsible for the hyperacute rejection in pig-to-human xenotransplantation,[27, 28] and, thus, there is a good demand for its large-scale production.[29] Synthesis of α-Gal epitope through a galactosidase-catalyzed reaction has previously been reported.[30] By taking advantage of the known Leloir's pathway, a biosynthetic pathway to Galα1,3Lac with recycling of UDP-Gal can be designed with five enzymes (Scheme 2). Galactokinase (GalK, EC 2.7.7.6) first converts galactose into galactose-1-phosphate with the consumption of one equivalent of phosphoenolpyruvate (PEP). Two enzymes, galactose-1-phosphate uridylyltransferase (GalT, EC 2.7.1.10) and glucose-1-phosphate uridylyltransferase (GalU, EC 2.7.1.9), catalyze galactose-1-phosphate and uridine 5'-triphosphate (UTP) to form UDP-Gal and pyrophosphate (PPi). This process involves glucose-1-phosphate and uridine 5'-diphosphoglucose (UDP-Glc) as intermediates. A galactosyltransferase will then transfer the galactosyl residue from UDP-Gal to lactose (the acceptor) to form the product. The resulting uridine 5'-diphosphate (UDP) can be phosphorylated to uridine 5'-triphosphate (UTP) by pyruvate kinase (PykF, EC 2.7.1.40) with the consumption of another equivalent of PEP. Overall, production of one equivalent of α-Gal epitope requires one equivalent each of galactose and lactose, and two equivalents of PEP (Scheme 2). In the in vitro system, PEP and/or adenosine 5'-triphosphate (ATP) have to be added to drive the glycosylation. However, in living bacterial cells, the high-energy phosphates may come from the normal cellular metabolism.

As reported herein, we assembled the genes of these five enzymes to form an artificial gene cluster on a single plasmid. Subsequent transformation of the plasmid into an *E. coli* strain affords a carbohydrate-producing bacterial strain—we call it a "superbug". Such a superbug can be fermented in large quantities and used for the synthesis of the desired oligosaccharide from inexpensive starting materials.

Overall:

$$\text{Gal} + \text{Lac} + 2\text{PEP} \rightleftharpoons \text{Gal}\alpha1,3\text{Lac} + 2\text{Pyruvate} + \text{PPi} \qquad (2)$$

Scheme 2. In vitro biosynthetic pathway of α-Gal epitope.

Results

Cloning, overexpression, and characterization of individual enzymes

In order to obtain individual enzyme activity, the pET15b vector was used to achieve individual expression of the enzymes GalK, GalT, GalU, PykF, and α1,3-galactosyltransferase (α1,3GT), all of which are involved in the α-Gal epitope biosynthetic pathway. The primers used in the cloning of individual genes are listed in Table 1. Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) indicated that each enzyme was overexpressed

Table 1. List of oligonucleotide primers used in the construction of individual genes in the pET15b vector and in the construction of the plasmid pLDR20-αKTUF for the α-Gal epitope producing superbug.

| Primer name	Primer sequence
For cloning individual enzyme genes in pET15b:	
galK-F	5'-GATCATATGAGTCTGAAAGAAAAAACAC-3'
galK-R	5'-CGCGGATCCTCAGCACTGTCCTGCTCCTTG-3'
galT-F	5'-GGATCCATATGACTAGTATGACGCAATTTAATCCC-3'
galT-R	5'-AGCGGATCCTTACACTCCGGATTCGCG-3'
galU-F	5'-GGATCCTCGAGATGCCTGCCATTAATACG-3'
galU-R	5'-CGCGGATCCACTAGTTTACTTCTTAATGCCCATCTC-3'
pykF-F	5'-GGATCCATATGAAAAAGACCAAAATTGTTTGCACC-3'
pykF-R	5'-CGCGGATCCACTAGTTTACAGGACGTGAACAGATGC-3'
For constructing multiple expression vector pLDR20-αKTUF:	
galU-F'	5'-CCGGATATCCCGCGGGTCGACAATAATTTTGTTTAACTTTAAGAAGG-3'
galU-R'	5'-GCATCGATGGTCTAGAGGATCCTTACTTCTTAATGCCCATCTC-3'
α1,3GT-F'	5'-GGATCCATATGACTAGTGATATCAATAATTTTGTTTAACTTTAAGAAGG-3'
α1,3GT-R'	5'-CCATCGATGTCGACCCGCGGTCAGACATTATTTCTAACCAC-3'
galKT-F'	5'-TCCCCGCGGCCCGGGAATAATTTTGTTTAACTTTAAGAAGG-3'
galKT-R'	5'-CGCGTCGACTCAGCACTGTCCTGCTCCTTG-3'

in the host strain with the target protein representing more than 70 % of the total soluble protein (data not shown). Consistently, 25, 100, 100, 50, and 10 U of GalK, GalT, GalU, PykF, and α1,3GT, respectively, were obtained from a 1 L expression fermentation as determined by activity assays carried out on enzymes purified with a nickel–nitrilotriacetic acid (Ni^{2+}-NTA) agarose affinity column. (See the Experimental Section for the definition of a Unit.)

Construction of the super plasmid and coexpression of enzymes

The successful construction of plasmid pLDR20-αKTUF for coexpression of the five enzymes is shown in Scheme 3. The primers used are listed in Table 1 and the detailed procedures were described in the Experimental Section. The resulting plasmid (Figure 1) contains a single λ P$_R$ promoter, a cl857 repressor gene, an ampicillin resistance gene (Amp^r), and a single T7 terminator. Each gene (except the *galK* gene, which has a natural ribosome binding sequence overlapped with the coding sequence of the upstream *galT* gene in the *gal* operon) was preceded by a Shine-Dalgarno sequence (*rbs*) for ribosome binding. This plasmid was transformed into a DH5α or NM522 *E. coli* strain to form α-Gal epitope producing cells DH5α (pLDR20-αKTUF) or NM522 (pLDR20-αKTUF). A *lacZ$^-$ E. coli* strain, such as DH5α or NM522, was used to eliminate the hydrolysis of lactose by the β-galactosidase that would be produced by other hosts, such as BL21(DE3). The target genes in pLDR20 were overexpressed by increasing the culture temperature from 30 to 40 °C. As shown by SDS-PAGE (Figure 2, lanes 2 – 4), coexpression had no adverse effect on the levels of soluble recombinant proteins.

Synthesis of α-Gal epitope with whole cells

The synthesis of Galα1,3Lac trisaccharide with NM522 (pLDR20-αKTUF) cells was analyzed under different conditions. Under the conditions listed in Table 2, run 1, the reaction produced a concentration of 16 mM of the final trisaccharide. About 66 % of the lactose was converted into the product. Further analysis showed that deviation from the above conditions, such as lowering the concentration of
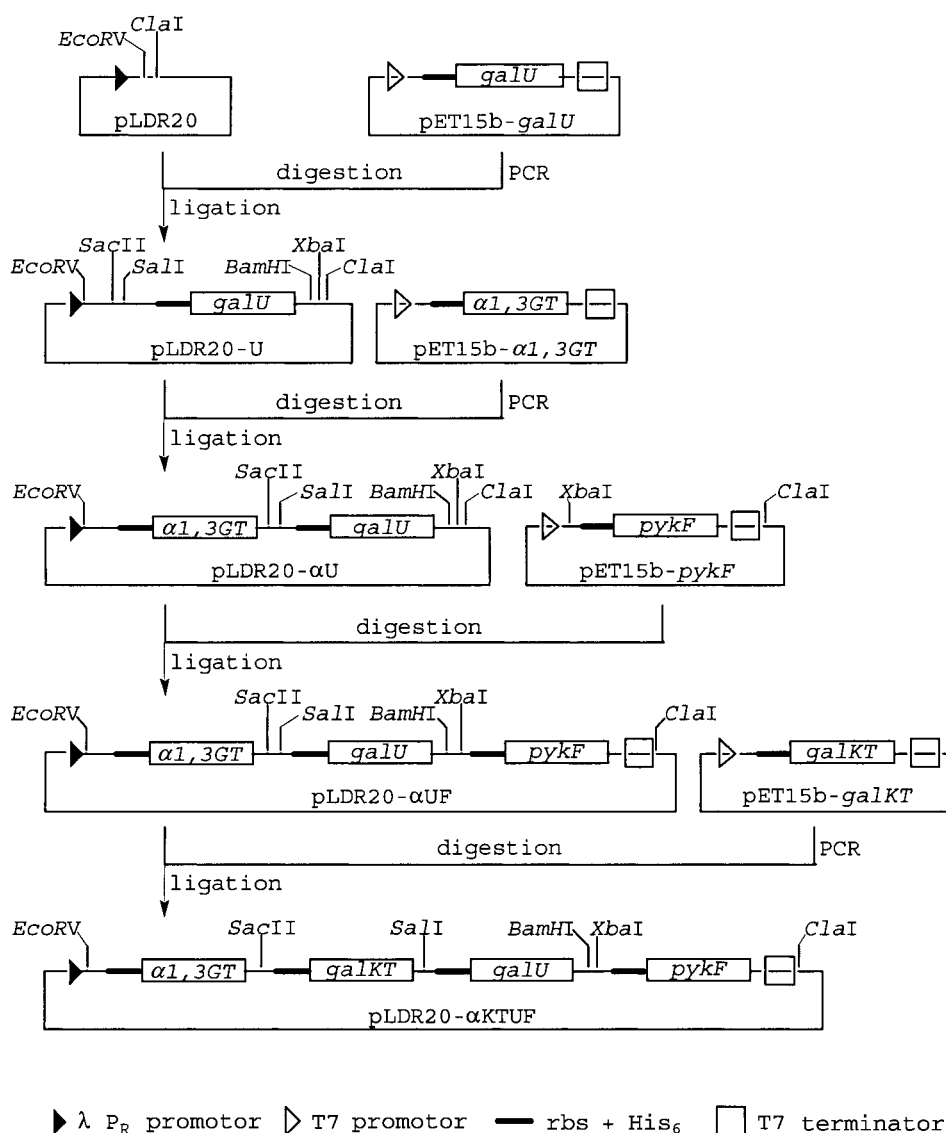
Scheme 3. *Construction of plasmid pLDR20-αKTUF. Digestion sites used for inserting subsequent genes have been highlighted.*
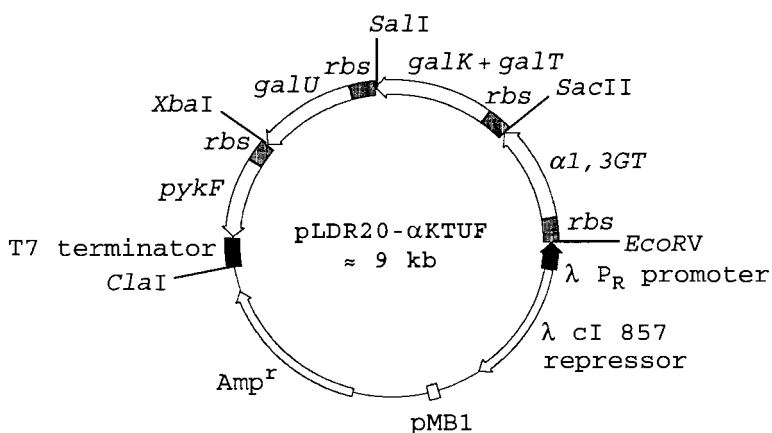
Figure 1. *Plasmid map of α-Gal epitope producing superbug harboring five genes encoding enzymes involved in the biosynthetic pathway of α-Gal oligosaccharides with UDP–Gal regeneration. Introduced restriction sites: EcoRV, SacII, SalI, XbaI, and ClaI.*

Table 2. *Production of Galα1,3Lac trisaccharide with the superbug under different conditions.*[a]

Run	Starting materials	Product concentration [mM]	Yield [%][b]
1	Gal (50 mM), Lac (25 mM), Glc (50 mM), PEP (5 mM), Glc-1-P (2 mM), UDP-Glc (2 mM), ATP (2 mM), MgCl₂ (10 mM), KCl (100 mM), MnCl₂ (10 mM), HEPES (50 mM)	16.5	66
2	conditions as in run 1 but with 0.5 mM UDP-Glc	13.5	54
3	conditions as in run 1 but with no Glc	13.3	53
4	conditions as in run 1 but with no PEP	13.5	54
5	conditions as in run 1 but with no ATP	10.0	40
6	economic conditions: Gal (40 mM), Lac (25 mM), Glc (20 mM), Glc-1-P (2 mM), UDP-Glc (0.5 mM), ATP (2 mM), MgCl₂ (10 mM), KCl (10 mM), MnCl₂ (10 mM)	10.8	43
7[c]	large-scale reaction: 10 L fermentation (65 g of cells), 500 mL reaction, other conditions as in run 6	14.2	57 (3.6 g)

[a] The reactions were performed with a 1 mL reaction volume. [b] The yield of product was determined by HPLC. [c] The large-scale reaction was performed in a total volume of 500 mL.
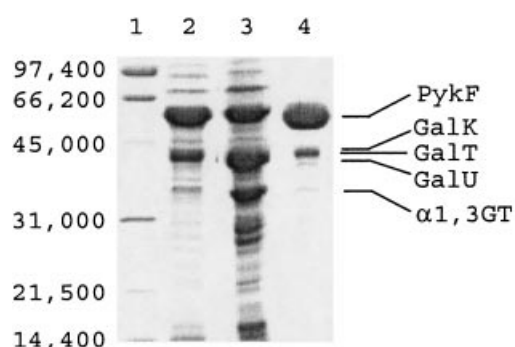

Figure 2. *SDS-PAGE showing proteins expressed in the α-Gal producing superbug. Lane 1: low molecular weight standards; lane 2: cell lysate of DH5α (pLDR20-αKTUF); lane 3: DH5α (pLDR20-αKTUF) whole-cell extract; lane 4: affinity-purified enzymes from the lysate of DH5α (pLDR20-αKTUF). GalK is present in lanes 2 and 3, but not in lane 4.*

Figure 3. *Time course of production of Galα1,3Lac trisaccharide with superbug NM522 (pLDR20-αKTUF). The reaction was performed in a 500 mL reaction volume and the product was monitored by HPLC.*

UDP-Glc from 2 to 0.5 mM, or eliminating Glc, PEP, or ATP, resulted in slightly lower yields. Since UDP-Glc is the most expensive component and an excess of the Gal and Glc starting materials will affect the downstream purification, the conditions for run 6 in Table 2 were the optimal compromise when both the yield and the cost were considered. For large-scale synthesis, 65 g of wet cells, produced from a 10 L fermentation, were used in a 500 mL reaction. A time course of the reaction (Figure 3) monitored by high-pressure liquid chromatography (HPLC) analysis with a refractive index (RI) detector indicated that the appearance of product reached a plateau after 20 hours at room temperature.

Discussion

This work successfully transferred an in vitro biosynthetic cycle into an oligosaccharide producing *E. coli* strain. A unique feature of this approach is the coexpression of five enzymes involved in the oligosaccharide biosynthetic pathway in one microbial strain to modify the host cells towards the accumulation of the desired product. After several failed attempts, we found that it was necessary to add a ribosomal binding sequence upstream of each component gene in the construction of the plasmid pLDR20-αKTUF to assure adequate translation of all the
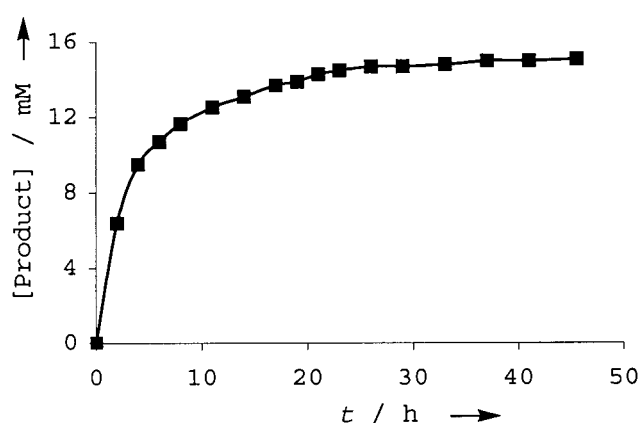
enzymes. However, since both *galK* and *galT* genes exist in the same *gal* operon, an additional ribosomal binding sequence is not necessary for the *galK* gene. The requirement of a ribosomal binding site for each gene was also exemplified in the work by Townsend and co-workers on the coexpression of three genes from *E. carotovora*. This was, to our knowledge, the only previous example of the use of the same strategy for multiple-gene coexpression.[31]

Another advantage of our strain is the use of internal sources of energy to drive the desired glycosylation cycle. The designed synthetic cycle (as shown in Scheme 2) requires two equivalents of PEP to form one equivalent of trisaccharide product. This requirement was also verified by using purified recombinant enzymes to conduct the glycosylation cycle (data not shown). However, when whole recombinant *E. coli* cells were used as catalysts, only catalytic amounts of ATP and UDP-Glc were needed. For the typical large-scale synthesis, 0.5 mM UDP-Glc was used to produce 14.2 mM trisaccharide. This indicates that the proposed cycle in Scheme 1 was repeated at least 30 times during the synthesis. Presumably the required high-energy molecules (such as PEP or ATP) were obtained through the glycolytic pathway or other metabolic pathways within *E. coli*. Due to the simplicity and efficiency of this superbug system, production of α-Gal epitope now becomes much more eco-

nomically feasible. Without extensive optimization, the current system produces 3 – 4 g of Galα1,3Lac trisaccharide from every 10 L fermentation. Based on the commercial prices for the chemical reagents and growth medium used, the material cost for the production is about $ 25 per gram of product. Larger-scale fermentation and production should most certainly reduce the costs further.

Compared to the currently available whole-cell approaches in oligosaccharide synthesis, the superbug technology presented herein obviously shows advantages. The Kyowa Hakko technology is efficient and cost effective in the production of sugar – nucleotides. However, this system involves several plasmids and multiple bacterial strains. In contrast, the superbug approach achieves the same goal with the use of only one plasmid and a single strain. All the enzymes essential for oligosaccharide synthesis, including the glycosyltransferases, and the sugar – nucleotide regeneration are in one *E. coli* strain. Thus, this approach avoids unnecessary transport of the intermediates in biosynthetic cycles from one strain to another for use in subsequent reactions. Instead of the *E. coli* only being modified by introduction of the glycosyltransferase genes, the superbug technology also overexpressed related enzymes essential for the regeneration of the sugar – nucleotide conjugate required by the glycosyltransferase. Combined with an effectual fermentation process, the superbug technology will be a more efficient method for large-scale production of complex carbohydrates.

It is important to point out that the superbug synthesis is different from well-known metabolic engineering efforts, such as those targeted at the polyketide syntheses[32, 33] or the production of aromatics from glucose.[34, 35] Those processes are generally one-step procedures, which generate the product during the fermentation of the engineered microorganisms. The superbug approach is a two-step process, the first step involving the growth of the engineered microorganisms and the second step involving their use as a source of catalysts (enzymes), energy, and cofactors to produce the desired compound. The two-step process results in a higher product concentration (14.2 mM in this case) in the reaction mixture than the conventional fermentative processes. Moreover, with the superbug in hand, different starting materials can be used in the second step to produce natural product derivatives. For the "superbug" that produces α-Gal epitope, we have found that both galactose derivatives and a variety of lactose derivatives could be used as starting materials to obtain α-Gal epitope analogues. Therefore, the superbug approach might be more versatile and closer to the conventional organic synthesis procedures.

Our experience indicates that monosaccharides, most disaccharides (for example, lactose) and trisaccharides are readily transported in and out of the cells. However, the transport of oligosaccharides larger than trisaccharide may pose a problem. Several solutions can be proposed: 1) permeabilizing reagents can be added to facilitate transport across the membrane; 2) a secretion mechanism can be applied to help the export of large oligosaccharide products, as is the case in hyaluronan synthesis; 3) cell-free in vitro synthesis with immobilized recombinant enzymes can be employed as we have recently demonstrated.[36]

Another potential difficulty in the application of the superbug approach is how to express all the enzymes in an active, soluble form. Although there is no general rule, several common guidelines can be followed; these include expressing the minimum catalytic domains, fusing with other proteins, using different vectors and strains, modification of several residues to help protein folding, changing codons to those suitable for the host strain,[37] and optimizing fermentation conditions such as temperature and agitation speed. In multiple expression systems such as ours, some enzymes may be expressed at a lower level and may limit the rate of oligosaccharide production. This problem can be solved by finding out the rate-limiting enzymatic step and then optimizing the cycle by incorporating multiple copies of the gene for this enzyme into the superbug plasmid or the chromosomal DNA of the host *E. coli* strain. Other methods such as replacing the ribosomal binding sites or adding different promoters in front of the genes in the plasmid can also be applied to control individual protein expression levels.[38]

In summary, this work demonstrates that an artificial gene cluster could be created by linking the genes for a natural product biosynthetic pathway. Transformation of such gene cluster into an *E. coli* provided a strain that can be used to efficiently generate the natural product. Specifically, in the field of carbohydrate synthesis, this approach provides a practical solution to the large-scale production of complex carbohydrate compounds to meet the increasing demand for biomedical research and pharmaceutical development.

Experimental Section

Bacterial strains and plasmids: Plasmid vector pLDR20 and *E. coli* K-12, substrain MG1655, were purchased from the American Tissue Culture Collection (ATCC). Plasmid vector pET15b and *E. coli* BL21(DE3) competent cells [F^- *ompT hsdS$_B$(r_B^- m$_B^-$) gal dcm (DE3)*] were from Novagen. Plasmid pET15b-α1,3GT was constructed as described previously.[28] *E. coli* DH5α competent cells [*lacZΔM15 hsdR recA*] were from Gibco. *E. coli* NM522 competent cells [*supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5(r_K^- m$_K^+$) (F'proABlacl^{q}ZΔM15)*] were from Stratagene. For chemical analysis, thin-layer chromatography (TLC) was conducted on Baker Si250F silica gel TLC plates with a fluorescent indicator. ^{1}H and ^{13}C NMR spectra were obtained on a Varian VXR400 NMR spectrometer. Mass spectra (Fast atom bombardment) were run at the mass spectrometry facility at Wayne State University.

Cloning, overexpression, and purification of individual enzymes: DNA manipulations were performed as described before.[28] The *galK* gene, *galT* gene, *galU* gene, or *pykF* gene were amplified by polymerase chain reaction (PCR) from the *E. coli* K-12 chromosome and inserted into the *Nde*I and *Bam*HI (*Xho*I and *Bam*HI for *galU*) restriction sites of the pET15b vector. The resulting plasmids pET15b-*galK*, pET15b-*galT*, pET15b-*galU*, and pET15b-*pykF* were transformed into the *E. coli* DH5α cloning host strain and then the BL21(DE3) expression strain. Selected clones were characterized by restriction mapping. The expression and purification of individual enzymes from the cell lysate was performed as described before.[39] Briefly, the overexpression was induced by 0.4 mM isopropyl-1-thio-β-D-galactos-pyranoside (IPTG) for 3 h at 37 °C in an incubator shaker. The cells were pelleted, resuspended, and then lysed enzymatically (by adding lysozyme). Unbroken cells and debris were pelleted by centrifugation

at 12 000 rpm for 20 min. The enzymes were purified by using a Ni^{2+}-NTA agarose affinity column which binds the N-terminal His 6-tag sequence in the recombinant proteins. After elution, the fractions containing the purified enzyme were combined and dialyzed for enzyme activity assays and enzymatic reactions.

Enzymatic assays: In this work, one unit of enzyme activity is defined as the amount of enzyme that produces 1 μmol of product per minute at 24 °C (37 °C for α1,3GT).

Galactokinase: The assay was performed at 24 °C for 30 min in a final volume of 100 μL in 2-[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulfonic acid (HEPES) buffer (pH 7.4; 100 mM) containing α-D-[6-^{3}H]galactose (0.5 mM, final specific activity of 1000 cpm nmole^{-1}) and ATP (50 mM). ATP was omitted for the blank experiment. The reaction was stopped by adding Dowex $1 \times 8 - 200$ chloride anion exchange resin suspended in water (0.8 mL). After centrifugation, supernatant (0.4 mL) was collected in a plastic vial and ScintiVerse BD (5 mL) was added. The radioactivity of the mixture was counted in a liquid scintillation counter.

Galactose-1-phosphate uridylyltransferase: This was a two-step assay. First, the GalT-catalyzed reaction was performed at 24 °C for 15 min in HEPES buffer (pH 7.4; 100 mM; final volume: 250 μL) containing 1.6 mM galactose-1-phosphate (Gal-1-P), 2.8 mM UDP-glucose, and enzyme solution (100 μL). A blank experiment was prepared with water replacing the Gal-1-P. The reaction was stopped by adding cold NaCl solution (0.5 mL, 0.15 M) and immediately transferring the tube to a boiling water bath for 5 min. After 15 min centrifugation at 14 000 g, the supernatant (0.2 mL) was subjected to the UDP-glucose assay in a cuvette with a total volume of 1 mL containing 0.03 M tris(hydroxymethyl) aminomethane (Tris)/acetate buffer (pH 8.7), 1.36 mM NAD, sample (0.2 mL) and UDP-glucose dehydrogenase (3.2 mU). The OD change at 340 nm was monitored by a UV spectrophotometer.

Glucose-1-phosphate uridyltransferase: The GalU-catalyzed reaction was carried out at 24 °C for 15 min in a final volume of 250 μL containing 1.6 mM glucose-1-phosphate (Glc-1-P), 2.8 mM UTP, 10 mM $MgCl_2$, and enzyme solution (100 μL). The blank experiment contained water instead of Glc-1-P. The reaction was stopped by adding cold NaCl solution (0.5 mL, 0.15 M) and immediately transferring the tube into a boiling water bath for 5 min. After centrifugation at 14 000 g for 15 min, the clear supernatant (0.2 mL) was subjected to the UDP-glucose assay by using UDP-glucose dehydrogenase as described for GalT.

Pyruvate kinase: The assay was carried out in a cuvette with a solution (1 mL) containing 0.1 M Tris/HCl buffer (pH 8.0), 0.5 mM ethylenediaminetetraacetic acid (EDTA), 0.1 M KCl, 10 mM $MgCl_2$, 0.2 mM NADH, 1.5 mM adenosine 5′-diphosphate (ADP), lactate dehydrogenase (60 mU), and 5 mM PEP. A blank experiment was carried out with water replacing the ADP. The reactions were performed at 24 °C, and the absorbance at 340 nm was monitored.

α1,3-galactosyltransferase: An assay utilizing UDP-[6-^{3}H]Gal and lactose was used as described previously.[28]

Construction of superbug: By using the preconstructed plasmids pET15b-*galK*, pET15b-*galT*, pET15b-*galU*, pET15b-α*1,3GT*, and pET15b-*pykF* as the PCR templates, the genes of the enzymes involved in the biosynthetic pathway of α-Gal were subcloned one by one into the pLDR20 vector with the ribosomal binding site and the His$_6$-tag encoding sequence preceding each gene (Scheme 3) to form the final plasmid pLDR20-αKTUF. This pLDR20 vector contains an ampicillin resistance gene, a P_R promoter, and a cl857 repressor gene. Since *galK* and *galT* exist in the same *gal* operon and close to each other, they were cloned together into the pET15b vector and

then into the pLDR20 vector. Briefly, the stepwise construction of plasmid pLDR20-αKTUF was as follows: First, the *galU* gene was PCR amplified from pET15b-*galU* with primers *galU*-F′ and *galU*-R′ which introduced the *EcoR*V, *Sac*II, and *Sal*I, or *Xba*I and *Cla*I restriction sites respectively. The PCR product was digested with *EcoR*V and *Cla*I and inserted into the multiple-cloning site of the pLDR20 vector previously cut with the same enzymes. Successful cloning was verified by restriction mapping and the expression of the GalU was confirmed by SDS-PAGE. Second, primers α*1,3GT*-F′ and α*1,3GT*-R′ were used to amplify the α*1,3GT* gene from plasmid pET15b-α1,3GT. The PCR product was digested and inserted into the *EcoR*V and *Sac*II restriction sites of the plasmid pLDR20-U to form plasmid pLDR20-αU. Third, the pET15b-*pykF* plasmid was digested with *Xba*I and *Cla*I, and the smaller fragment containing the *pykF* gene, ribosomal binding site, sequence for the N-terminal His 6-tag, and T7 terminator was purified and inserted into the *Xba*I and *Cla*I restriction sites of the plasmid pLDR20-αU to form plasmid pLDR20-αUF. Finally, the *galK* and *galT* genes were amplified from plasmid pET15b-*galKT* (constructed by inserting the gene sequence encoding both *galK* and *galT* into a pET15b vector between the *Nde*I and *BamH*I restriction sites) by using the primers *galKT*-F′ and *galKT*-R′. The PCR product was digested and inserted into the *Sac*II and *Sal*I restriction sites of plasmid pLDR20-αUF to form plasmid pLDR20-αKTUF. The plasmid pLDR20-αKTUF was transformed into DH5α or NM522 competent cells to make DH5α (pLDR20-αKTUF) and NM522 (pLDR20-αKTUF) α-Gal epitope producing cells.

Synthesis of α-Gal epitope with whole cells: α-Gal epitope producing superbug NM522 (pLDR20-αKTUF) was grown in 4 L shake flasks. The expression of the target genes in the superbug was initiated by increasing the temperature from 30 to 40 °C. After shaking at 40 °C for 3 – 3.5 h, the cells were separated from the medium by centrifugation (4000 g for 20 min) and suspended in 20 mM Tris/HCl buffer (pH 8.5; 100 mL) containing 1% Triton X-100. For better results, the cell suspension was freeze-thawed twice before being applied in the reaction. For small-scale analysis, the reaction was performed with cells (0.14 g, wet weight) in a 1 mL reaction volume containing Gal (50 mM), Lac (25 mM), Glc (50 mM), PEP (5 mM), Glc-1-P (2 mM), UDP-Glc (2 mM), ATP (2 mM), $MgCl_2$ (10 mM), KCl (100 mM), $MnCl_2$ (10 mM), and HEPES (50 mM, pH 7.4). The reaction was carried out at room temperature and the formation of the trisaccharide product was monitored by HPLC with a MICROSORB 100Å amino column with CH_3CN/H_2O (65:35) as the eluent. To optimize the conditions, multiple 1 mL reactions were set up with different starting material compositions.

Large-scale synthesis of Galα1,3Lac trisaccharide: Large-scale production consists of two steps, culturing the superbug cells and then trisaccharide production catalyzed by the cells. NM522 (pLDR20-αKTUF) cells were first grown at 30 °C in a 10 L fermentor, then enzyme expression was induced by increasing the temperature to 40 °C for 3 h. The cells were separated from the media by centrifugation. The cell pellet (65 g, wet weight) was stored at −20 °C and frozen/thawed twice before use in the reaction. For a typical synthesis reaction, the NM522 (pLDR20-αKTUF) cells (65 g in Tris/HCl buffer (pH 8.5; 20 mM; 250 mL) containing 1% Triton X-100) were added to HEPES buffer (pH 7.4; 0.1 M; 250 mL) containing Lac (4.28 g, 12.5 mmol), Gal (3.60 g, 20 mmol), Glc (1.80 g, 10 mmol), ATP (0.55 g, 1.0 mmol), UDP-Glc (0.15 g, 0.25 mmol), Glc-1-P (0.31 g, 1.0 mmol), $MnCl_2$ (20 mM), $MgCl_2$ (20 mM), and KCl (20 mM), to bring the total reaction volume to 500 mL. The reaction was stirred at 24 °C for 36 h and was monitored by thin-layer chromatographic analysis (*i*PrOH/NH_4OH/H_2O (7:3:2)). HPLC analysis indicated that the concentration of Galα1,3Lac was around 14 mM (7.2 g L^{-1}, 56% yield based on acceptor Lac). After the reaction, the cells were removed by

centrifugation and the remaining lactose was hydrolyzed to the monosaccharides by incubating the supernatant with β-galactosidase (0.2 g, Sigma) for 10 h. The mixture was then poured onto a column packed with graphitized carbon (Supelco). The column was washed with water. The trisaccharide product (3.6 g) was eluted with water containing 10% acetonitrile and the solvent was removed under reduced pressure. The product was characterized by NMR spectroscopy and mass spectrometry. ^{1}H NMR (500 MHz, D$_2$O; selected peaks): $\delta = 5.04$ (d, $J = 3.6$ Hz, 0.4 H), 4.96 (d, $J = 3.6$ Hz, 1 H), 4.48 (d, $J = 7.6$ Hz, 0.6 H), 4.36 (d, $J = 7.6$ Hz, 1 H), 3.99 (m, 2 H), 3.38 – 3.84 (m, 12 H), 3.11 (t, $J = 8.1$ Hz, 1 H); ^{13}C NMR (125 MHz, D$_2$O): $\delta = 102.93$, 95.88, 95.52, 91.93, 78.74, 78.61, 77.26, 75.16, 74.88, 74.54, 73.88, 71.60, 71.22, 70.92, 70.17, 69.69, 69.39, 69.22, 68.32, 64.91, 61.14, 61.03, 60.25, 60.12; HR-MS (FAB): calcd for C$_{18}$H$_{32}$O$_{16}$ [M+Na]: 527.1588, found: 527.1582.

[1] J. R. Jacobsen, C. R. Hutchinson, D. E. Cane, C. Khosla, *Science* **1997**, *277*, 367 – 369.

[2] L. E. Quadri, J. Sello, T. A. Keating, P. H. Weinreb, C. T. Walsh, *Chem. Biol.* **1998**, *5*, 631 – 645.

[3] Y. Xue, L. Zhao, H. W. Liu, D. H. A. Sherman, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12 111 – 12 116.

[4] R. McDaniel, A. Thamchaipenet, C. Gustafsson, H. Fu, M. Betlach, G. Ashley, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1846 – 1851.

[5] L. Tang, S. Shah, L. Chung, J. Carney, L. Katz, C. Khosla, B. Julien, *Science* **2000**, *287*, 640 – 642.

[6] C.-H. Wong, S. L. Haynie, G. M. Whitesides, *J. Org. Chem.* **1982**, *47*, 5416 – 5418.

[7] C.-H. Wong, G. M. Whitesides, *Enzymes in Synthetic Organic Chemistry*, Vol. 12, Elsevier Science, Amsterdam, **1994**, pp. 252 – 297.

[8] C.-H. Wong, R. L. Halcomb, Y. Ichikawa, T. Kajimoto, *Angew. Chem.* **1995**, *107*, 569 – 593; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 521 – 546.

[9] M. Gilbert, R. Bayer, A. M. Cunningham, S. DeFrees, Y. Gao, D. C. Watson, N. M. Young, W. W. Wakarchuk, *Nat. Biotechnol.* **1998**, *16*, 769 – 772.

[10] M. M. Palcic, *Curr. Opin. Biotechnol.* **1999**, *10*, 616 – 624.

[11] J. C. Paulson, *Trends Biochem. Sci.* **1989**, *14*, 272 – 276.

[12] X. Chen, P. Kowal, P. G. Wang, *Curr. Opin. Drug Discovery Dev.* **2000**, *3*, 756 – 763.

[13] G. F. Herrmann, P. Wang, G.-J. Shen, C.-H. Wong, *Angew. Chem.* **1994**, *106*, 1346 – 1347; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 1241 – 1242.

[14] X. Chen, W. Zhang, J. Wang, J. Fang, P. G. Wang, *Biotechnol. Prog.* **2000**, *16*, 595 – 599.

[15] G. F. Herrmann, L. Elling, C. H. Krezdorn, R. Kleene, E. G. Berger, C. Wandrey, *Bioorg. Med. Chem. Lett.* **1995**, *5*, 673 – 676.

[16] E. Samain, S. Drouillard, A. Heyraud, H. Driguez, R. A. Geremia, *Carbohydr. Res.* **1997**, *302*, 35 – 42.

[17] E. Samain, V. Chazalet, R. A. Geremia, *J. Biotechnol.* **1999**, *72*, 33 – 47.

[18] E. Bettler, E. Samain, V. Chazalet, C. Bosso, A. Heyraud, D. H. Joziasse, W. W. Wakarchuk, A. Imberty, R. A. Geremia, *Glycoconjugate J.* **1999**, *16*, 205 – 212.

[19] E. Samain, B. Priem, WO Patent 00104341, **2000**.

[20] T. Endo, S. Koizumi, K. Tabata, S. Kakita, A. Ozaki, *Carbohydr. Res.* **2001**, *330*, 439 – 443.

[21] T. Endo, S. Koizumi, *Curr. Opin. Struct. Biol.* **2000**, *10*, 536 – 541.

[22] K. Tabata, S. Koizumi, T. Endo, A. Ozaki, *Biotechnol. Lett.* **2000**, *22*, 479 – 483.

[23] T. Endo, S. Koizumi, K. Tabata, A. Ozaki, *Appl. Microbiol. Biotechnol.* **2000**, *53*, 257 – 261.

[24] T. Endo, S. Koizumi, K. Tabata, S. Kakita, A. Ozaki, *Carbohydr. Res.* **1999**, *316*, 179 – 183.

[25] S. Koizumi, T. Endo, K. Tabata, A. Ozaki, *Nat. Biotechnol.* **1998**, *16*, 847 – 850.

[26] U. Galili, J. L. Avila, *α-Gal and Anti-Gal*, Kluwer Academic/Plenum, New York, **1999**.

[27] A. H. Good, D. K. C. Cooper, A. J. Malcolm, R. M. Ippolito, E. Koren, F. A. Neethling, Y. Ye, N. Zuhidi, L. R. Lamontagne, *Transplant. Proc.* **1992**, *24*, 559 – 562.

[28] W. Parker, S. Saadi, S. S. Lin, Z. E. Holzknecht, M. Bustos, J. L. Platt, *Immunol. Today* **1996**, *17*, 373 – 378.

[29] X. Chen, P. Andreana, P. G. Wang, *Curr. Opin. Chem. Biol.* **1999**, *3*, 650 – 658.

[30] G. Vic, M. Scigelova, J. J. Hastings, O. W. Howarth, D. H. Crout, *Chem. Commun.* **1996**, *12*, 1473 – 1474.

[31] R.-F. Li, A. Stapon, J. T. Blanchfield, C. A. Townsend, *J. Am. Chem. Soc.* **2000**, *122*, 9296 – 9297.

[32] A. F. Marsden, B. Wilkinson, J. Cortes, N. J. Dunster, J. Staunton, P. F. Leadlay, *Science* **1998**, *279*, 199 – 202.

[33] J. T. Kealey, L. Liu, D. V. Santi, M. C. Betlach, P. J. Barr, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 505 – 509.

[34] J. W. Frost, K. M. Draths, *Annu. Rev. Microbiol.* **1995**, *49*, 557 – 579.

[35] R. Patnaik, R. G. Spitzer, J. C. Liao, *Biotechnol. Bioeng.* **1995**, *46*, 361 – 370.

[36] X. Chen, J. Fang, J. Zhang, Z. Liu, P. Andreana, J. Shao, P. Kowal, P. G. Wang, *J. Am. Chem. Soc.* **2001**, *123*, 2081 – 2082.

[37] S. J. Higgins, B. D. Hames, *Protein expression: a practical approach*, Oxford University Press, New York, **1999**.

[38] J. L. Cleland, C. S. Craik, *Protein engineering: principles and practice*, Wiley-Liss, New York, **1996**.

[39] J. W. Fang, J. Li, X. Chen, Y. N. Zhang, J. Q. Wang, Z. M. Guo, W. Zhang, L. B. Yu, K. Brew, P. G. Wang, *J. Am. Chem. Soc.* **1998**, *120*, 6635 – 6638.