重组DNA技术用于生产维生素C重要前体——2-酮基-L-古龙酸的新进展

尹 光 琳

（上海交通大学生物科学与技术系，上海）

重组DNA技术在生产实践中的价值，一方面涉及到使用便于迅速培养的微生物或动植物宿主细胞来制备外源蛋白质或多肽；另一方面，通过基因的分离、鉴别、重组、转移和表达等定向培育，从而有效地制取为人类需要的有价值的产物[1,2]。最近，美国的科学工作者对于维生素C的重要前体——2-酮基-L-古龙酸（以下简称2-KLG）的制备使用了建造的“代谢工程菌”重组体进行了成功的尝试。他们分离了棒状杆菌（*Corynebacterium* sp.）中的2,5-二酮基-D-葡萄糖酸（以下简称2,5-DKG）还原酶基因的克隆，建造了表达这一基因的载体，把它转化到一株草生欧文氏菌（*Erwinia herbicola*）中，扩展了D-葡萄糖的代谢途径，从而进一步发酵就可以得到维生素C的重要前体——2-KLG[3]。

维生素C生产的历史背景

维生素C（即L-抗坏血酸）是人体不可缺少的营养物质，在医药、食品、饮料和饲料等方面都有广泛的用途。目前全世界的维生素C工业年产值高达6亿美元。随着市场销售量不断增加，各国对于生产工艺的改革都非常重视。

1933年，Reichstein等人[4]用化学合成法制取维生素C获得成功，至今大多数国家仍一直采用经过修改的这个方法（通称“莱氏法老工艺”）进行生产。这一生产过程长而反应剧烈的路线包括微生物一步发酵和一系列化学步骤，生成的中间体是2-KLG，经过简单的酸或碱催化的环化作用（内酯化）转化生成维生素C（图1）。六十年代初期，美国[5]、法国[6]、日本[7]和西

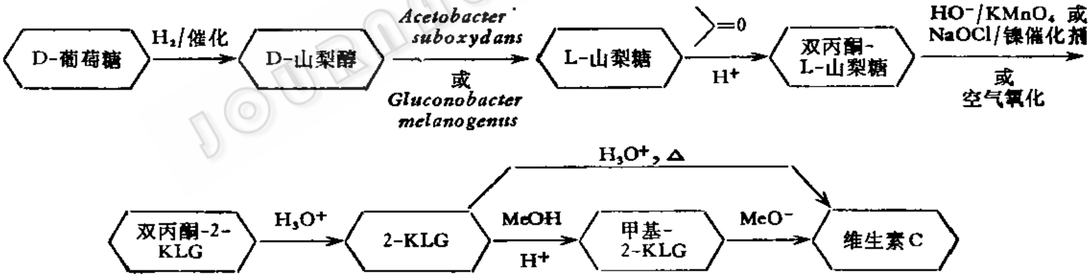

图1 莱氏法合成维生素C流程图

德等国相继进行了简化和缩短莱氏法反应路线的研究，注意点都集中在微生物发酵上，使用D-山梨醇或L-山梨糖作为底物来直接生成2-KLG。但因其中水平最高者发酵率才46%左右，故都只停留在实验室阶段并未进行生产性扩大试验。七十年代初期，由中国科学院微生物研究所和北京制药厂合作[8]，首先筛选到能利用L-山梨糖产生2-KLG的优良混合菌株，以后经过国内许多工厂不断实践改进和提高，自1974年6月中试鉴定会后已逐步在全国推广。这就是常说的二步发酵法（通称“新工艺”）。它是以一步发酵所得的L-山梨糖作为底物，用一株氧化葡萄糖酸杆菌（*Gluconobacter oxydans* Subsp.）（产酸菌，通称“小菌”）和一株假单胞杆菌（*Pseudomonas* sp.）或芽孢杆菌（*Bacillus* sp.）（搭配用伴生菌，通称“大菌”）组成的混合菌株继续发酵，即可直接氧化生成2-KLG（图2）。二步发酵法的主要优点是能把莱氏法中的三步化学反应简化成微生物发酵，从而大大简化了生产工序，革除和减少对有毒及易爆的化学药品的消耗并降低了原料成本。目前，国内生产维生素C的工厂中采用二步发酵法新工艺的占三分之二以上，菌株性能一直很稳定，发酵率为80%左右，维生素C总收率在47%以上。二步发酵法的研究在国际上来说，我国至今仍处于领先地位。

从七十年代中期以后，日本园田高安等人进行了直接从D-葡萄糖发酵经过中间体2.5-DKG产生2-

© 中国科学院微生物研究所期刊联合编辑部 http://journals.im.ac.cn

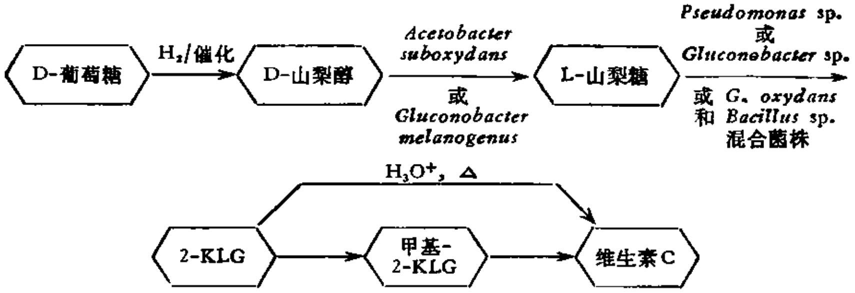

图 2 二步发酵法生成维生素C流程图

KLG 的研究[9,10]。他们用一株欧文氏菌 (*Erwinia* sp.) 的突变株和一株棒状杆菌 (*Corynebacterium* sp.) 进行串联发酵 (a tandem fermentation process) 完成了从 D-葡萄糖到 2-KLG 的转化。这一串联发酵过程简化了从 D-葡萄糖到最后合成维生素C的反应路线,然而它还是应用了两株菌,分两步发酵完成(图 3 上部分),有待于进一步改进[11,12]。

用重组 DNA 技术构建"代谢工程菌"生产 2-KLG

为了实现对上述两个菌株两步发酵的进一步简化和更有效地生产 2-KLG,美国 Genentech 公司 Anderson 等人[3] 首先鉴定了棒状杆菌中 2, 5-DKG 还原酶的作用,证实了这一还原酶承担了从 2,5-DKG到 2-KLG 这一代谢步骤的生物还原作用。他们将此还原酶的基因克隆出来并构建了表达载体,把它转化到草生欧文氏菌中表达,从而实现了从 D-葡萄糖到 2-KLG 的一步发酵,图 3 下部分表示了用单个重组体菌株进行一步发酵的可能性,这就是利用重组 DNA 技术建造新的"代谢工程菌"产生 2-KLG 的设想。

(一) 探针的合成及棒状杆菌中 2,5-DKG 还原酶基因的分离和鉴定

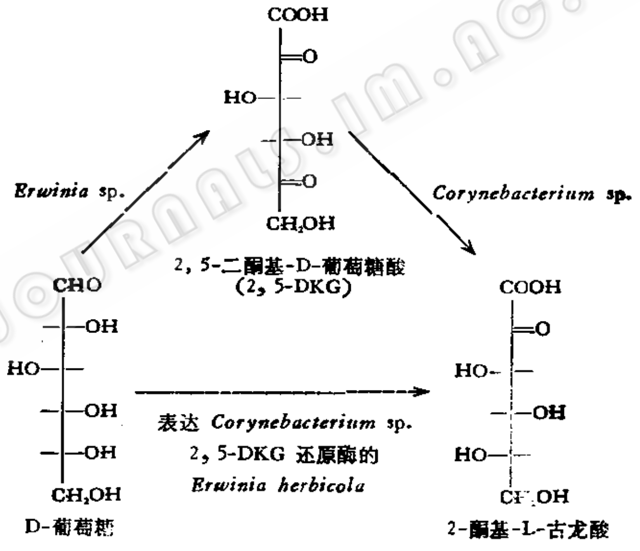

图 3 从 D-葡萄糖直接发酵生成 2-酮基-L-古龙酸示意图

为了从棒状杆菌中分离获得 2,5-DKG 还原酶基因,必须制备合适的多核苷酸探针,它的顺序和该基因的部分编码区的顺序应该有最大的均一性[13]。因此,首先纯化和鉴定了 2,5-DKG 还原酶,并测定了它的-NH_2 端的 40 个氨基酸顺序,由此设定了合成两个探针的核苷酸顺序。根据测定结果,探针 A 为 CTCCATCCCGCAGCTGGGGCTACGGCGTGTTCAAGGTGCCGCCG,它是 2,5-DKG 还原酶中第 11 至 25 位氨基酸编码区的同义链 (Same Sense);探针 B 为 GGCCTCCTCCACGGCGCGCTGGGCGTCGGCCGGCGGCACCTTG,它是 2,5-DKG 还原酶中第 21 至 35 位氨基酸编码区的反义链(图 4)。利用这两个探针,对棒状杆菌作了基因组的 Southern 印迹法测定,确定了两个探针全都和 2.2 Kb 的内切酶 BamHI 片段进行杂交,因而在制备规模上分离得到了棒状杆菌的 2.0—2.5Kb 的 BamHI 片段,把它克隆到质粒载体 pBR 322 中,将得到的部分基因库采用同样的探针作菌落原位杂交,结果证明所筛选得到的克隆虽然包含了绝大部分的 2,5-DKG 还原酶基因,但缺少了这个基因的 3'-端部分编码区和终止区(图 5 核苷酸顺序中第 874 位以后的区域),进而又

© 中国科学院微生物研究所期刊联合编辑部 http://journals.im.ac.cn

```
                                                                                       Met Thr Val
1    CCATGGACGGCAGCCCGAACGATCCCGACACCATCGTGCTCGACCCGCACGCCTGGAGGGCGCGGACCGCGCCTACCCTGGAAGAC ATG ACA GTT
     Nco I

     Pro Ser Ile Val Leu Asn Asp Gly Asn Ser Ile Pro Gln Leu Gly Tyr Gly Val Phe Lys Val Pro Pro Ala Asp
96   CCC AGC ATC GTG CTC AAC GAC GGC AAT TCC ATT CCC CAG CTC GGG TAC GGC GTC TTC AAG GTG CCG CCG GCG GAC

     Thr Glu Arg Ala Val Glu Glu Ala Leu Glu Val Gly Tyr Arg His Ile Asp Thr Ala Ala Ile Tyr Gly Asn Glu
171  ACC CAG CGC GCC GTC GAG GAA GCG CTC GAA GTC GGC TAC CGG CAC ATC GAC ACC GCG GCG ATC TAC GGA AAC GAA

     Glu Gly Val Gly Ala Ala Ile Ala Ala Ser Gly Ile Ala Arg Asp Asp Leu Phe Ile Thr Thr Lys Leu Trp Asn
246  GAA GGC GTC GGC GCC GCG ATC GCG GCG AGC GGC ATC GCG CGC GAC GAC CTG TTC ATC ACG ACG AAG CTC TGG AAC

     Asp Arg His Asp Gly Asp Glu Pro Ala Ala Ala Ile Ala Glu Ser Leu Ala Lys Leu Ala Leu Asp Gln Val Asp
321  GAT CGC CAC GAC GGC GAT GAG CCC GCT GCA GCG ATC GCC GAG AGC CTC GCG AAG CTG GCA CTC GAT CAG GTC GAC
                                         Pst I

     Leu Tyr Leu Val His Trp Pro Thr Pro Ala Ala Asp Asn Tyr Val His Ala Trp Glu Lys Met Ile Glu Leu Arg
396  CTG TAC CTC GTG CAC TGG CCG ACG CCC GCC GCC GAC AAC TAC GTG CAC GCG TGG GAG AAG ATG ATC GAG CTT CGC

     Ala Ala Gly Leu Thr Arg Ser Ile Gly Val Ser Asn His Leu Val Pro His Leu Glu Arg Ile Val Ala Ala Thr
471  GCA GCC GGT CTC ACC CGC AGC ATC GGC GTC TCG AAC CAC CTC GTG CCG CAC CTC GAG CGC ATC GTC GCC GCC ACC

     Gly Val Val Pro Ala Val Asn Gln Ile Glu Leu His Pro Ala Tyr Gln Gln Arg Glu Ile Thr Asp Trp Ala Ala
546  GGC GTC GTG CCG GCG GTG AAC CAG ATC GAG CTG CAC CCC GCC TAC CAG CAG CGC GAG ATC ACC GAC TGG GCC GCC

     Ala His Asp Val Lys Ile Glu Ser Trp Gly Pro Leu Gly Gln Gly Lys Tyr Asp Leu Phe Gly Ala Glu Pro Val
621  GCC CAC GAC GTG AAG ATC GAA TCG TGG GGG CCG CTC GGT CAG GGC AAG TAC GAC CTC TTC GGC GCC GAG CCC GTC

     Thr Ala Ala Ala Ala Ala His Gly Lys Thr Pro Ala Gln Ala Val Leu Arg Trp His Leu Gln Lys Gly Phe Val
696  ACT GCG GCT GCC GCC GCC CAC GGC AAG ACC CCG GCG CAG GCC GTG CTC CGT TGG CAC CTG CAG AAG GGT TTC GTG
                                                                                  Pst I

     Val Phe Pro Lys Ser Val Arg Arg Glu Arg Leu Glu Glu Asn Leu Asp Val Phe Asp Phe Asp Leu Thr Asp Thr
771  GTC TTC CCG AAG TCG GTC CGC CGC GAG CGC CTC GAA GAG AAC CTC GAC GTG TTC GAC TTC GAC CTC ACC GAC ACC

     Glu Ile Ala Ala Ile Asp Ala Met Asp Pro Gly Asp Gly Ser Gly Arg Val Ser Ala His Pro Asp Glu Val Asp
846  GAG ATC GCC GCG ATC GAC GCG ATG GAT CCG GGC GAC GGT TCG GGT CGC GTG AGC GCA CAC CCC GAT GAG GTC GAC
                                     Bam HI
         ⌐—Δ270                                       |→ URF
921  TGA CCCCGCCGAACACCCGGAGGCCACGGGCGCAGGACTAGCCTGGGCTCGTGGCATCCCGGCTCTCCGAAGCGACCAGCCCCTACCTGCGCGCCCAC

1019 GCCGACAACCCCGTCGCATGGTGGCCGTGGGGCGAGGCGGCCTTCGCAGAAGCCCGTCGCCGCGACGTGCCCGTGATGGTCTCGATCGGCTACTCGACAT

1119 GCCACTGGTGCCACGTGATGGCACGGGAGAGCTTCGAGGATGCCGCGGTCGCCGCCGACCTCGACGCAGGGTTCGTCGCCGTCAAGGTCGACCGCGAGGA

1219 GCATCCCGAGGTCGACGCGGCCTACATGGCGGCCGCCGCCGCATTCACGCAGAACCTCGGCTGGCCGCTCACCGTCTTCGTCACACCCGCGGGCGACCG

1319 TTCTTCGCGGGAACCTACTTCCCACCCGAACCGCGCGGCGGACTGCCTGCGTTCCGGCAGGTGCTCGCCGCCGTCGACGAGGCCTGGACCGAACGCCGCG

1419 ACCAGATCGAGAGCACAGGCGGCGCGATCGTGGATGCCCTCGCCGAGGTGCGGGGCGTCGCCGGTGCCGAGGCGTACGCGCTGCCGTCCGTCGACGACCT

1519 CGCCGCTGCCGCGACGGCGCTCGCCGCCCGTGAAGACACGGAGTTCGGCGGGTTCGGCGCAGCGGGCGGATCGCTCGAACAGCCCAAGTTCCCCGTCGCG

1619 ACCGCACTGCGGTTCCTGCAG
              Pst I
```

图 4 棒状杆菌 (*Corynebacterium* sp.) 2,5-DKG 还原酶基因顺序

利用了已得到的克隆中 0.12Kb 的片段(第 753—874 位)作为探针，分离获得了另一个含有 2,5-DKG 还原酶基因的 3′-端编码区和终止区的 0.9Kb 的内切酶 PstI 片段(第 753—1639 位)的克隆，进一步作成亚克隆。其插入片段共有 1639 个核苷酸长，它包含了棒状杆菌中 2,5-DKG 还原酶基因的全部顺序。这个基因的编码中含有氨基酸 278 个残基的蛋白质。从 DNA 顺序推断出的氨基酸顺序和 2,5-DKG 还原酶蛋白的 $-NH_2$ 端顺序及其它几个胰蛋白酶酶解的多肽顺序是十分吻合的。由于这个 2,5-DKG 还原酶基因在内切酶 NcoI 位点和 ATG 起始密码之间(第 1—89 位)没有起动子顺序可以用来识别，就在 NcoI 位点"上游"("upstream") 插入了乳糖操纵子 (lac) 或色氨酸操纵子 (trp) 的起动子。从测定结果可以看到，2,5-DKG 还原酶基因不能在大肠杆菌 (*E. coli*) MM 294 中直接被表达。

(二) 2,5-DKG 还原酶表达载体的构建

为了使 2,5-DKG 还原酶基因得到表达，对载体作了适当改造(图 5)。第一步是为了把色氨酸操纵子起动子放在适当的位置上，必须把 ATG 起始密码前面的核苷酸(图 4 中所示第 1—87 位)突变删除 (缺失 deletion)，把整个插入片段转入载体 M13 mp 9 中，构成 mit 12 质粒。mit 12 的单链 DNA 用缺失引物 (ACGGCCAGTCAATTCTATGACAGTTCCGAGC) 和 Alu I 片段作退火，再用 DNA 聚合酶大片段 (Klenow) 合成互补链，用 DNA 连接酶接上，最后转化到大肠杆菌 (*Ecoli*) JM 101 菌中并筛选出缺失突变体 (图 5 中的 mit 12△)。第二步即在 mit 12△ 的插入片段前面 (即图 4 中 ATG 起始密码"上游") 接上质粒 pHGH 207-1 ptrp △R15 (Amp^R) 的色氨酸操纵子起动子和核糖体结合位点，并把它们插入质粒 pBR322/Hind III、Pst I 中，构建得到的表达载体为 pmit 12△/trpI(Amp^R)。

© 中国科学院微生物研究所期刊联合编辑部 http://journals.im.ac.cn

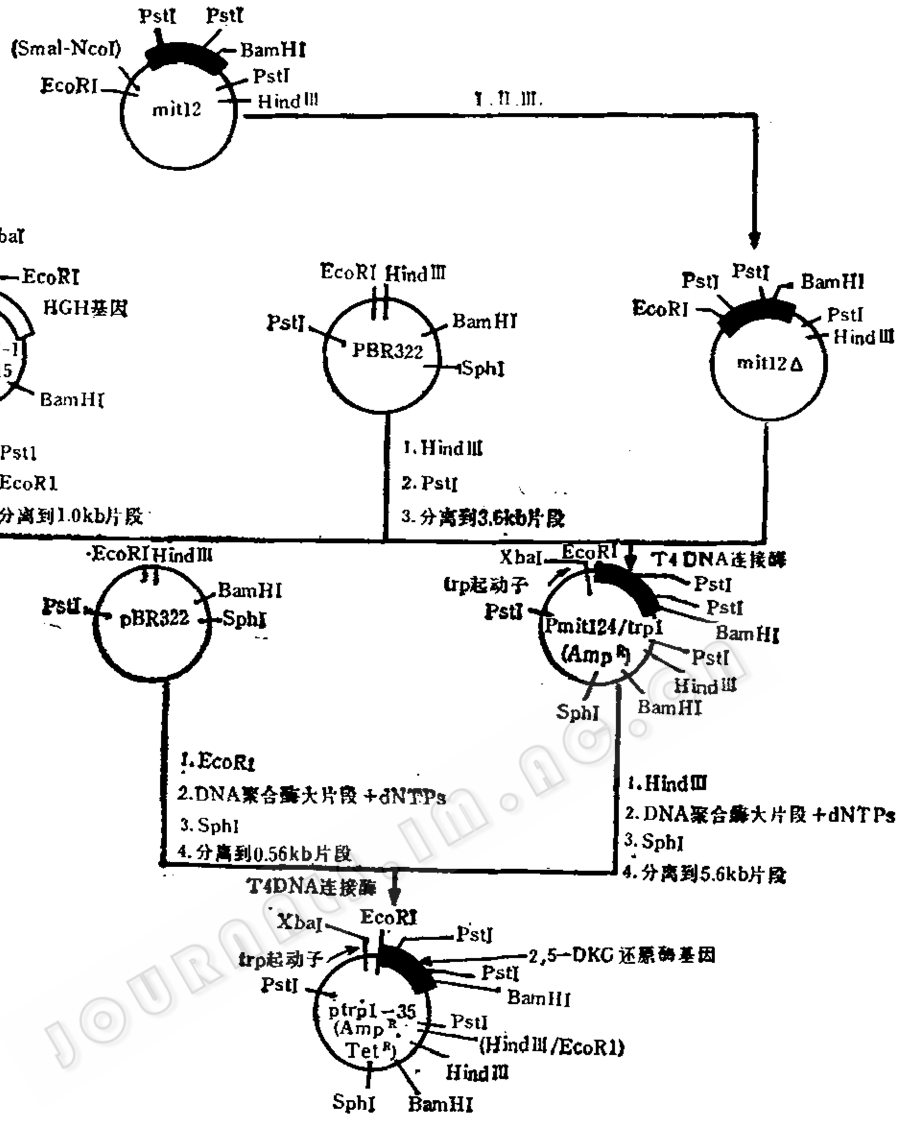

图 5 2,5-二酮基-D-葡萄糖酸（2,5-DKG）还原酶表达载体的构建

I. 删除(缺失)引物及 Alu I 片段和 mit 12 ssDNA 的退火 II. DNA 聚合酶大片段+dNTPs + T4DNA 连接酶 III. 转化 JM 101 和删除突变体的筛选

实验证实了这一表达载体在大肠杆菌（*E. coli*）MM 294 中得到表达，在它的细胞可溶性抽提液中具有 2,5-DKG 还原酶的活性。在用 SDS-聚丙烯酰胺凝胶电泳后作免疫印迹法[14]的结果显示出有与棒状杆菌中 2,5-DKG 还原酶相同迁移率的一条蛋白质区带。

然而，考虑到这个表达质粒（pmit 12Δ/trpI）是对氨苄青霉素有抗性而对四环素是敏感的，如果要实现对欧文氏菌的转化，必须把它变成对四环素有抗性才行，因为欧文氏菌也是对氨苄青霉素有抗性而对四环素是敏感的。因此，对最后一步表达质粒的改建是在 pmit 12Δ/trpI 中插入质粒 pBR 322 中的四环素起动子，得到的最终表达载体 ptrp I-35（Amp^R Tet^R）在经过转化到草生欧文氏菌中后，即可用四环素抗性作为选择标记来筛选所期望的转化子。

（三）表达的效果

在转化 ptrp I-35 的草生欧文氏菌的可溶性细胞抽提液中，其 2,5-DKG 还原酶的活力是 55.4n mol/分/A_{280}，相比之下，用 pBR 322 作转化的对照抽提液的酶活力只有 1.2n mol/分/A_{280}。在 SDS-聚丙烯酰胺凝胶电泳中，ptrpI-35 转化抽提液中具有一条和 2,5-DKG 还原酶相同迁移率的显著区带，但在质粒 pBR 322 的转化细胞中则没有这条区带，用免疫印迹法（Immunoblot）[14] 来测定可以证实这一区带就是 2,5-DKG 还原酶。

© 中国科学院微生物研究所期刊联合编辑部 http://journals.im.ac.cn

利用 pup 1-35 转化的草生欧文氏菌（*E. herbicola*）在含有D-葡萄糖的培养基（ATCC 1038 培养基）中生长达到饱和后，取出离心并将细菌细胞再悬浮到含有2%甘油的新鲜1038培养基中，进一步培养后将培养液作高效液相层析（HPLC），发现2-酮基-L-古龙酸的浓度为1g/l；并未检测到还有D-葡萄糖和其他有机酸存在，而将质粒 pBR 322 转化到欧文氏菌中的对照菌液检测却并未见到有2-酮基-L-古龙酸存在。产物经气相色谱、质谱分析、^{1}H 和 ^{13}C 核磁共振谱测定以及旋光和熔点分析，可证明确实是2-酮基-L-古龙酸。

结　束　语

上面介绍了 Anderson 等人的工作，使用重组 DNA 技术，把两个极不相同的微生物中的代谢途径合并到一个微生物中，从而用这个“代谢工程菌”进行一步发酵就可完成从D-葡萄糖到2-酮基-L-古龙酸的生物转换作用。从表面上看，这只是一个简单的表达载体的转化，但在细胞内实际情况更要复杂一些。因为从D-葡萄糖进行一系列氧化作用转化到2,5-DKG的脱氢酶[15]是固定在细胞膜上的周膜酶（periplasmic enzyme），而2,5-DKG还原酶则是可溶性的，存在于细胞浆内，完成这两步反应并使得所生成的产物排出细胞外，就需要进行对这些小分子物质的胞内外迁移[16]。只有具备这样的条件后，重组体的草生欧文氏菌（recombinant *E. herbicola*）才能有效地从D-葡萄糖生产维生素C的重要中间体——2-酮基-L-古龙酸。

重组 DNA 技术的研究为应用微生物学增加了一个新领域。Genentech 公司 Anderson 等人构建“代谢工程菌”来生产2-酮基-L-古龙酸这一非常出色研究的成功，标志着利用遗传工程进行工业规模大量生产化学制品的时代已经到来。美国著名的 Pfizer 公司、Lubrizol Enterprises 公司已和 Genentech 公司签订协议，将共同开发这一新技术[17]。这对于我国开展遗传工程的研究是一个启迪。我们深信，随着重组DNA这一新兴技术的深入研究，将为生物工程的实际应用开发更为广阔的未来。

参　考　文　献

[1] Boyer, H. W. et al.: *Sci. News*, 112(20): 310, 1977.
[2] Ensley, B. D. et al.: *Science*, 222: 167, 1983.
[3] Anderson, S. et al.: *Science*, 230: 144, 1985.
[4] Reichstein, T. et al.: Helv. Chim. Acta., 16: 1019, 1933.
[5] Huang, H. T.: U. S. Patent, 3, 043, 749, 1962.
[6] Fr. Patent, 1, 376, 741, C 07c — C 12k, 1964.
[7] 望月一男等：特许公报，昭 37—43475，1962。
[8] 尹光琳等：微生物学报，20：246，1980。
[9] 阋田高安等：特许公报，昭 54—19468，1979。
[10] Kita, D. A. et al.: U. S. Patent, 4,316,960, 1982.
[11] Sonoyama, T. et al.: *Appl. Environ. Microbiol.*, 43: 1064, 1982.
[12] Sonoyama, T. et al.: Eur. Patent, 0,088,409 1983.
[13] Anderson, S. et al.: *Proc. Natl. Acad. Sci. U. S. A.*, 80: 6838, 1983.
[14] Renart, J. et al.: ibid., 76: 3116, 1979.
[15] Ameyama, M. et al.: *Agric. Biol. Chem.*, 45: 851, 1981.
[16] Ameyama, M. et al.: ibid., 45: 1079, 1984.
[17] Chemical Eng. News, 63(46): 6, 1985.

© 中国科学院微生物研究所期刊联合编辑部 http://journals.im.ac.cn