

# Cloning gibberellin dioxygenase genes from pumpkin endosperm by heterologous expression of enzyme activities in *Escherichia coli*

(2-oxoglutarate-dependent dioxygenases/*in vitro* translation/functional, enzyme activity, heterologous expression/fusion protein/cDNA cloning)

THEO LANGE\*

Pflanzenphysiologisches Institut und Botanischer Garten der Universität Göttingen, Untere Karspüle 2, D-37073 Göttingen, Germany

Communicated by Jake MacMillan, University of Bristol, Bristol, United Kingdom, April 3, 1997 (received for review December 3, 1996)

**ABSTRACT** Gibberellin (GA) plant hormones are biosynthesized via complex pathways, the final steps of which are catalyzed by 2-oxoglutarate-dependent dioxygenases. Here, the cloning of two such enzymes, the GA 7-oxidase and the GA 20-oxidase, is reported using a novel approach, namely, by screening for GA dioxygenase activities expressed as T7 gene 10 fusion proteins in recombinant *Escherichia coli*. *In vitro* translation products of mRNA from endosperm of immature pumpkin seeds contained three GA dioxygenase activities, including 7-oxidase, 20-oxidase, and 3 $\beta$ -hydroxylase. A cDNA expression library was prepared from the endosperm mRNA in  $\lambda$ MOSElox. An aliquot of the amplified library was converted to plasmids *in vivo* and used for transformation of *E. coli* BL21(DE3), which thereafter expressed recombinant fusion proteins containing 7-oxidase, 20-oxidase, and 3 $\beta$ -hydroxylase activities. By screening for specific GA dioxygenase expression, clones harboring 7-oxidase and 20-oxidase cDNA were isolated. The ORF of the 7-oxidase cDNA is 945 bp long, encoding for 314 amino acid residues with a calculated  $M_r$  of 35,712 and pI of 5.7. Recombinant GA 7-oxidase oxidizes GA<sub>12</sub>-aldehyde to GA<sub>12</sub> and GA<sub>14</sub>-aldehyde to GA<sub>14</sub>. Evidence was obtained for further metabolism of GA<sub>12</sub> by the 7-oxidase to four products, two of which are monohydroxylated GA<sub>12</sub>. The ORF of the 20-oxidase is—apart from seven changes, resulting in four amino acid substitutions—identical to the 20-oxidase cDNA previously cloned from pumpkin cotyledon mRNA; both 20-oxidases have the same catalytic properties.

Gibberellins (GAs) control many important processes in the life cycle of higher plants, such as germination, growth, flower induction, fruit set, and fruit development (1). Their biosynthesis has been extensively studied in immature pumpkin seeds (2–4), which are a rich source of GA biosynthetic enzymes and their encoding mRNAs (5–7). The biosynthetic pathway proceeds from the common intermediate, GA<sub>12</sub> aldehyde, as illustrated in Fig. 1. First, the oxidation at C-7 converts the aldehyde to the carboxylic acid to give GA<sub>12</sub>. Then, sequential oxidation at C-20 eventually produces two kinds of products: a tricarboxylic acid (GA<sub>25</sub>) or, with loss of C-20 and lactonization, a C<sub>19</sub>-GA (GA<sub>9</sub>). 3 $\beta$ -Hydroxylation produces a parallel series leading to the corresponding C-20 carboxylic acid (GA<sub>13</sub>) and C<sub>19</sub>-lactone (GA<sub>4</sub>). GA<sub>13</sub> is 2 $\beta$ -hydroxylated to GA<sub>43</sub>, which is a major constituent of pumpkin endosperm. C<sub>19</sub>-GAs are considered to be plant hormones, because they are active in bioassays, whereas C-20 tricarboxylic acids often are regarded as byproducts. However, the significance of high concentrations of GAs in immature seeds is to a large extent

unknown and, thus, it cannot be judged which GA derivatives are active here and which are not.

Recently, the structural gene for a GA 20-oxidase was cloned from pumpkin seed cotyledons by immunoscreening (7). The cloning of another 20-oxidase gene from pumpkin endosperm was reported, but details of this work were not documented (8). Subsequently, many GA 20-oxidase genes have been cloned from various plant sources (8–15) using PCR-based cloning procedures. All 20-oxidases are shown to be multifunctional, catalyzing the whole series of oxidations at C-20, and are responsible for the partition between C<sub>20</sub>- and C<sub>19</sub>-GAs, which is much in favor of C-20 carboxylic acids with the pumpkin enzyme, whereas GA 20-oxidases from other sources produce mainly C<sub>19</sub>-GAs. Recently the cloning of a putative GA 3 $\beta$ -hydroxylase gene has been reported by a T-DNA tagging strategy (16).

Cloning of genes often is hindered by the limited availability of specific probes. Here, a novel cloning strategy is described, which circumvents this difficulty by assaying pools of recombinant *E. coli* cells for heterologous expression of enzyme activity to identify clones containing GA 7-oxidase and 20-oxidase cDNAs.

## MATERIALS AND METHODS

**Plant Material, Poly(A)<sup>+</sup> RNA Isolation, and *in Vitro* Translation.** Endosperm of seeds of *Cucurbita maxima* L. cv. “Riesenmelone, gelb genetzt” (van-Waveren, Göttingen, Germany) at different developmental stages was used. To identify the most advantageous starting material, poly(A)<sup>+</sup> RNA was extracted from endosperm of immature seeds without developed cotyledons, with cotyledons of 20%, and with cotyledons of 40% the length of the seed lumen (17) and translated *in vitro* using rabbit reticulocyte lysate (7). The products were directly assayed for GA dioxygenase activities by incubation with [<sup>14</sup>C]GA<sub>12</sub>-aldehyde, [<sup>14</sup>C]GA<sub>12</sub>, or [<sup>14</sup>C]GA<sub>25</sub> (250 Bq each, 40 pmol) and extraction as described (7). These precursors would assure detection of GA 7-oxidase, 20-oxidase, and 3 $\beta$ -hydroxylase, respectively. HPLC analysis was performed as described below.

**$\lambda$ -MOSElox cDNA Phage Library Construction.** Poly(A)<sup>+</sup> RNA (5  $\mu$ g) from immature seeds without developed cotyledons was used for the preparation of an oligo(dT)-primed cDNA expression library in  $\lambda$ MOSElox using commercial kits (Amersham). The library consisted of  $3 \times 10^6$  plaque-forming units, 30% of which contained inserts >350 bp as shown by

Abbreviations: GA, gibberellin; cfu, colony-forming units; RACE, rapid amplification of cDNA ends.

Data deposition: The sequences of the cDNA encoding the GA 7-oxidase and GA 20-oxidase have been deposited in the GenBank database (accession nos. U61386 and U61385, respectively).

\*e-mail: tlange@gwdg.de.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1997 by The National Academy of Sciences 0027-8424/97/946553-6\$2.00/0

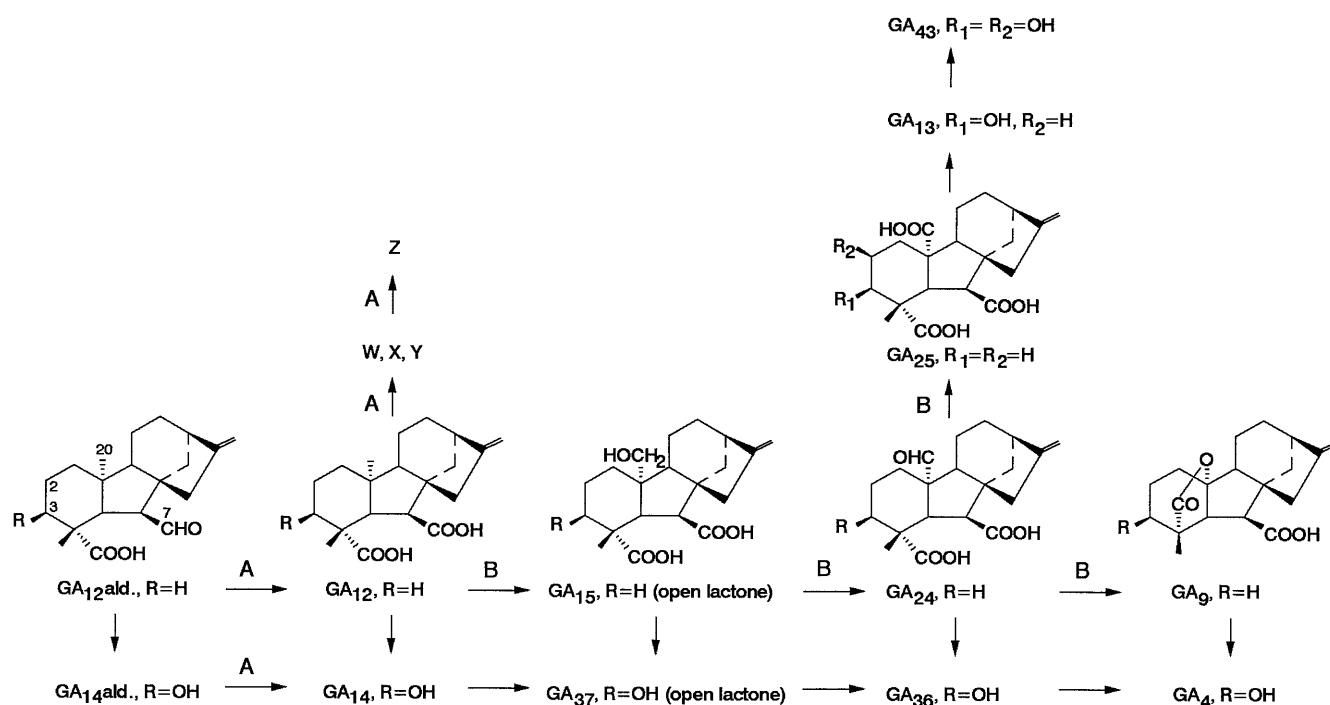


FIG. 1. GA biosynthetic pathways to  $C_{20}$ -tricarboxylic acids ( $GA_{25}$ ,  $GA_{13}$ ,  $GA_{43}$ ) and  $C_{19}$ -GAs ( $GA_9$ ,  $GA_4$ ) as found in pumpkin endosperm (2, 3) and reactions shown to be catalyzed by recombinant 7-oxidase (A) and 20-oxidase (B).

agarose gel electrophoresis of PCR products using  $\lambda$ MOSElox specific primers. One million plaque-forming units of the library were amplified in *E. coli* strain ER1647, according to the manufacturer's instructions.

**pMOSElox cDNA Plasmid Library Construction.** *E. coli* BM25.8 [ $2.5 \times 10^8$  colony forming units (cfu)] were infected by recombinant phage from the amplified  $\lambda$ MOSElox cDNA-library ( $10^9$  plaque-forming units) to convert pMOSElox plasmids by cre recombinase-mediated excision according to the manufacturer's instructions (Amersham). Approximately 292,000 cfu containing plasmids were obtained and amplified by growing in 50 ml of L-broth, supplemented with carbenicillin (100  $\mu$ g/ml) for 16 hr at 37°C with shaking to ensure aeration. The pMOSElox plasmids were purified using Plasmid Midi Kit (Qiagen) and used for transformation of *E. coli* BL21(DE3) according to ref. 18.

**Screening of pMOSElox cDNA Library by Heterologous Expression of GA 7-Oxidase and 20-Oxidase Activity.** For primary screening transformed *E. coli* BL21(DE3) were subdivided into 49 tubes, each containing  $\approx 790$  cfu in 1 ml of L-broth, supplemented with carbenicillin (50  $\mu$ g/ml), and grown for 16 hr at 37°C with shaking. The cultures were arranged in a  $7 \times 7$  grid, and 100- $\mu$ l aliquots from each of seven tubes across columns, and from each of seven tubes down rows, were pooled in a similar manner as described for PCR screening (19). The pooled bacteria were used for inoculation of 50 ml of  $2 \times$  yeast tryptone medium (YT), supplemented with carbenicillin (50  $\mu$ g/ml), and grown at 30°C with shaking until  $OD_{600}$  reached 0.8–1.0. Recombinant protein expression was induced by isopropyl  $\beta$ -D-thiogalactoside, and GA dioxygenase activities were determined directly in cell lysates as described later. Only protein from pooled bacteria of row E and from pooled bacteria of column 3 converted [ $^{14}$ C]GA<sub>12</sub>-aldehyde. Products were [ $^{14}$ C]GA<sub>12</sub> (3%) and [ $^{14}$ C]GA<sub>15</sub> (7%), as identified by HPLC, indicating that tube E3 contained GA dioxygenase encoding cDNA. Bacteria from tube E3 were subdivided for secondary screening into 64 tubes, each containing  $\approx 37$  cfu and reamplified and rescreened using an  $8 \times 8$  grid as described above for the primary screen. Bacteria of one tube (designated E4) converted [ $^{14}$ C]GA<sub>12</sub>-aldehyde to [ $^{14}$ C]GA<sub>12</sub>

and those of another one (C4) converted [ $^{14}$ C]GA<sub>12</sub> through to [ $^{14}$ C]GA<sub>25</sub>. Bacteria of both tubes were plated on separate L-broth plates, containing carbenicillin (50  $\mu$ g/ml), at  $\approx 100$  cfu per plate and grown for 22 hr at 37°C. Single colonies were transferred into 100  $\mu$ l of L-broth medium, containing carbenicillin (50  $\mu$ g/ml) and grown for 16 hr at 37°C with shaking. Rather than subjecting each clone to the laborious workup and assay, a third screening was performed. Ten microliters from each of 40 isolated and amplified clones originating from tube E4 was placed in an  $8 \times 5$  grid as described for the primary screen, and the pooled samples were amplified in culture volumes of 10 ml. Forty clones originating from tube C4 were screened exactly the same way. One clone of the first group was shown to oxidize [ $^{14}$ C]GA<sub>12</sub>aldehyde to [ $^{14}$ C]GA<sub>12</sub> (designated D4) and one clone of the second group was shown to convert [ $^{14}$ C]GA<sub>12</sub> to [ $^{14}$ C]GA<sub>25</sub>, via [ $^{14}$ C]GA<sub>15</sub> and [ $^{14}$ C]GA<sub>24</sub> (designated E5). The sizes of their cDNA inserts were determined by PCR. Both clones were used for heterologous expression studies and for DNA sequence analyses.

**Heterologous Expression of GA Dioxygenase Activities in *E. coli* and GA Dioxygenase Assay.** Recombinant *E. coli* BL21(DE3), harboring isolated clones D4 or E5 were grown in 50 ml of L-broth, supplemented with carbenicillin (50  $\mu$ g/ml), for 16 hr at 37° with shaking. Five milliliters of the cultures was inoculated into 250 ml of  $2 \times$  YT, containing carbenicillin (50  $\mu$ g/ml), at 30°C with shaking until  $OD_{600}$  reached 0.6–0.7. Expression was induced by the addition of isopropyl  $\beta$ -D-thiogalactoside to 1 mM, and, after further incubation at 30°C for 2 hr, cultures were transferred to 50-ml tubes and harvested by centrifugation at  $5,000 \times g$  for 5 min at 4°C. Pellets were resuspended in lysis buffer (2.5 ml, containing 200 mM Tris-HCl, pH 7.4, 1 mg/ml lysozyme, and 10 mM DTT), incubated on ice for 30 min, and freeze-thawed four times. The lysates were supplemented with 1  $\mu$ l of DNaseI (50 units, Sigma), incubated for 5 min at room temperature, and centrifuged at  $53,000 \times g$  for 30 min at 4°C. The supernatant ( $\approx 2.5$  ml) was stored at  $-80^\circ\text{C}$ . Culture volumes smaller than 250 ml, as used for screening the pMOSElox cDNA library, were treated by reducing the amount of all components accordingly. Cell lysates (70  $\mu$ l) were incubated at 30°C for 16 hr with [ $^{14}$ C]GA substrates (2  $\mu$ l, 250 Bq) in the

presence of cofactor mixture (30  $\mu$ l, containing 400 mM 2-oxoglutarate, 400 mM ascorbate, 2 mM FeSO<sub>4</sub>, and 4 mg/ml catalase). Variations in culture volumes, lysis buffer, and incubation conditions are specified for particular experiments. The substrates [<sup>14</sup>C]GA<sub>12</sub>-aldehyde (5.93  $\times$  10<sup>12</sup> Bq/mol) and [<sup>14</sup>C]GA<sub>12</sub> (5.92  $\times$  10<sup>12</sup> Bq/mol) were prepared as described (20). Some of the [<sup>14</sup>C]GA<sub>12</sub>-aldehyde (16 kBq) was converted to [<sup>14</sup>C]GA<sub>14</sub>-aldehyde (5.29  $\times$  10<sup>12</sup> Bq/mol) by partially purified GA 3 $\beta$ -hydroxylase (5). Incubation products were extracted and separated by HPLC with on-line radiocounting (20) using a H<sub>2</sub>O-methanol gradient as described (5). More than 95% of the radioactivity originally added as substrate was recovered. Retention times for the <sup>14</sup>C-labeled GAs were 28 min, 18 sec for GA<sub>12</sub>-aldehyde, 27 min, 12 sec for GA<sub>12</sub>, 22 min, 10 sec for GA<sub>14</sub>-aldehyde, 20 min, 8 sec for GA<sub>14</sub>, 18 min, 18 sec for compound W, 16 min, 16 sec for compound X, 12 min, 20 sec for compound Y, and 7 min, 48 sec for compound Z. Radioactive fractions were dried, derivatized, and analyzed by combined GC-MS as described (21). Mass spectra were acquired from 17 to 26 min after injection at an electron energy of 70 eV and a current of 600  $\mu$ A from 100 to 600 atomic mass units at 0.8 sec per scan.

**DNA Sequence Analysis.** The cDNA insert of clone D4 was excised using *Eco*RI and further digested with *Pst*I. The insert of clone E5 was released and digested with *Eco*RI. 5' Rapid amplification of cDNA ends (RACE) was performed for the clone D4 using the 5'/3' RACE Kit (Boehringer). Undigested inserts, digested fragments, and RACE products were subcloned using appropriate restriction sites of pBlueScript SK(-) (Stratagene), and transformants were selected with *E. coli* XL1blue. Plasmid DNA was isolated from single transformants using a Plasmid Midi Kit (Qiagen) and sequenced by the dideoxynucleotide chain-termination method from the -29 universal and T7-sequencing primers, 5'-labeled with IR-41 fluorescent dye (MWG Biotec, Ebersberg, Germany), using the Thermo Sequenase cycle sequencing kit, with 7-deaza-dGTP (Amersham) and a LiCor 4000L automated DNA sequencer (MWG). Homology searches of the GenBank, Protein DataBank, SwissProt, and Protein Information Resource data bank using the program BLAST and multiple sequence alignments using CLUSTAL W1.6 were carried out via the internet server at <http://kiwi.imgen.bcm.tmc.edu:8088/search-launcher/launcher.html>.

## RESULTS

**In Vitro Translation of Poly(A)<sup>+</sup> RNA.** *In vitro* translation products of poly(A)<sup>+</sup>RNA from pumpkin endosperm catalyzed the oxidation at C-7 of [<sup>14</sup>C]GA<sub>12</sub>-aldehyde to [<sup>14</sup>C]GA<sub>12</sub>, the successive oxidation at C-20 of [<sup>14</sup>C]GA<sub>12</sub> to [<sup>14</sup>C]GA<sub>15</sub>, [<sup>14</sup>C]GA<sub>24</sub> and [<sup>14</sup>C]GA<sub>25</sub>, and the hydroxylation at the 3 $\beta$ -position of [<sup>14</sup>C]GA<sub>15</sub>, [<sup>14</sup>C]GA<sub>24</sub> and GA<sub>25</sub> to [<sup>14</sup>C]GA<sub>37</sub>,

[<sup>14</sup>C]GA<sub>36</sub> and [<sup>14</sup>C]GA<sub>13</sub>, respectively (Table 1, Fig. 1). Furthermore a small proportion of the 2 $\beta$ ,3 $\beta$ -dihydroxylated product [<sup>14</sup>C]GA<sub>43</sub> was obtained in incubations with [<sup>14</sup>C]GA<sub>25</sub> indicating limited 2 $\beta$ -hydroxylation. Dioxygenase activity was highest in translation products of mRNA from very young endosperm and decreased rapidly with mRNA from older tissues (Table 1).

**Isolation of cDNA Clones.** For cloning of the GA 7-oxidase gene neither antibody nor nucleic acid sequence probes were available. However, both GA 7-oxidase and 20-oxidase activities were detectable in lysates from *E. coli* BL21(DE3) infected with the amplified  $\lambda$ -MOSElox cDNA library derived from pumpkin endosperm poly(A)<sup>+</sup> RNA (data not shown). Cell lysates prepared from a pool of 8,400 cfu of the amplified pMOSElox library converted [<sup>14</sup>C]GA<sub>12</sub> to [<sup>14</sup>C]GA<sub>15</sub> and [<sup>14</sup>C]GA<sub>37</sub> as identified by HPLC, when they were incubated with the appropriate cofactors (Fig. 2). Thus, both 20-oxidase and 3 $\beta$ -hydroxylase activities were present in these cell lysates. Cell lysates prepared from another pool, containing 59,000 cfu, converted GA<sub>12</sub>-aldehyde to GA<sub>15</sub> indicating that GA 7-oxidase and 20-oxidase activities were present in this portion of the library (Fig. 2). By expression of enzyme activity first in smaller pools, and finally in randomly picked clones, two clones were isolated, one encoding a 7-oxidase (D4) and one encoding a 20-oxidase (E5). The sizes of the cDNA inserts were  $\approx$ 1,150 bp and  $\approx$ 1,450 bp, respectively.

**Heterologous Expression in *E. coli*.** The catalytic properties of the recombinant 7-oxidase and 20-oxidase were investigated by expression of the respective cDNA molecules in pMOSElox in *E. coli* BL21 (DE3) and incubation of cell lysates with <sup>14</sup>C-labeled GA precursors. The results are summarized in Table 2. Recombinant 7-oxidase catalyzed oxidation at C-7 of [<sup>14</sup>C]GA<sub>12</sub>-aldehyde and [<sup>14</sup>C]GA<sub>14</sub>-aldehyde to [<sup>14</sup>C]GA<sub>12</sub> and [<sup>14</sup>C]GA<sub>14</sub>, respectively. [<sup>14</sup>C]GA<sub>12</sub>-aldehyde and [<sup>14</sup>C]GA<sub>12</sub> also were oxidized to four other products, W, X, Y, and Z. Compound X also was converted to product Z. The conversion of GA<sub>12</sub>-aldehyde to GA<sub>12</sub> and GA<sub>14</sub>-aldehyde to GA<sub>14</sub> (data not shown) and formation of compound X (Table 2) was more efficient at low pH and increased at higher lysate volumes at least up to 70  $\mu$ l per 100  $\mu$ l of standard incubation. Full-scan mass spectra of methyl ester trimethylsilyl ether-derivatives of compounds W and X revealed an *M*<sup>+</sup> of 448 for both compounds (Table 2), indicating that they are monohydroxylated GA<sub>12</sub> derivatives. To judge from its mass spectrum, compound W is probably 12 $\alpha$ - or 12 $\beta$ -hydroxy-GA<sub>12</sub>. The mass spectrum of compound X has no similarity to any known GA. Mass spectra of compounds Y and Z could not be determined because of the low yield of these products in incubations. No conversion was obtained of the <sup>14</sup>C-labeled substrates GA<sub>15</sub>, GA<sub>24</sub>, GA<sub>25</sub>, GA<sub>9</sub>, GA<sub>53</sub>, GA<sub>17</sub>, GA<sub>20</sub>, and GA<sub>1</sub>, and there was

Table 1. Metabolism of [<sup>14</sup>C]GA<sub>12</sub>-aldehyde, [<sup>14</sup>C]GA<sub>12</sub>, and [<sup>14</sup>C]GA<sub>25</sub> by *in vitro* translation products of poly(A)<sup>+</sup> RNA from endosperm of immature seeds of *Cucurbita maxima*, without developed cotyledons (0%), with cotyledons of 20%, and with cotyledons of 40% the length of the seed lumen (17)

Length of cotyledon, % of seed lumen	Substrate	Products, % of recovered radioactivity						
		GA <sub>12</sub> -ald.	GA <sub>12</sub>	GA <sub>15</sub> , GA <sub>24</sub> , GA <sub>25</sub>	GA <sub>37</sub> , GA <sub>36</sub> , GA <sub>13</sub>	GA <sub>43</sub>		
0	GA <sub>12</sub> -ald.	9.0*	2.6	68.8	19.4	0		
	GA <sub>12</sub>		3.9*	87.0	9.1	0		
	GA <sub>25</sub>			74.9*	23.0	2.1		
20	GA <sub>12</sub> -ald.	84.6*	2.3	9.3	3.8	0		
	GA <sub>12</sub>		67.2*	32.8	0	0		
	GA <sub>25</sub>			100*	0	0		
40	GA <sub>12</sub> -ald.	93.5*	6.5	0	0	0		
	GA <sub>12</sub>		100*	0	0	0		
	GA <sub>25</sub>			100*	0	0		

Products were identified by HPLC, and the sums of radioactivity found in peaks for non-3 $\beta$ -hydroxy GAs (GA<sub>15</sub>, GA<sub>24</sub>, and GA<sub>25</sub>) and of 3 $\beta$ -hydroxy GAs (GA<sub>37</sub>, GA<sub>36</sub>, and GA<sub>13</sub>) are given.

\*Unmetabolized substrate.

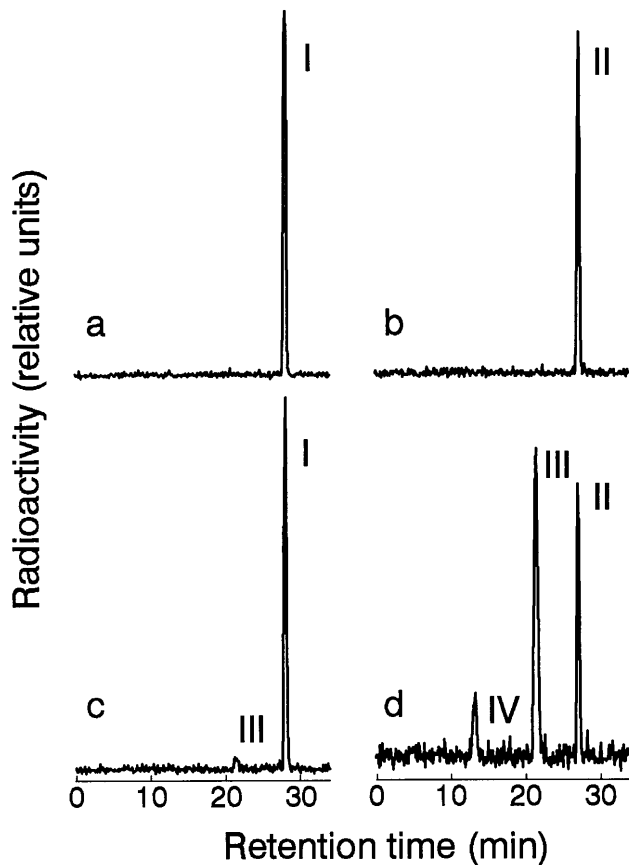


FIG. 2. GA 7-oxidase, 20-oxidase, and  $3\beta$ -hydroxylase activity in *E. coli* BL21(DE) harboring pMOSElox cDNA expression library. [ $^{14}\text{C}$ ]GA<sub>12</sub>-aldehyde (a, c) and [ $^{14}\text{C}$ ]GA<sub>12</sub> (b, d) were incubated with lysates prepared from bacteria with pMOSElox plasmids without insert (a, b) or with 59,000 (c) or 8,400 (d) clones of the pMOSElox cDNA library derived from pumpkin endosperm poly(A)<sup>+</sup> RNA. On the basis of HPLC retention times peaks correspond to GA<sub>12</sub>-aldehyde (I), GA<sub>12</sub> (II), GA<sub>15</sub> (III), and GA<sub>37</sub> (IV).

no expression of 7-oxidase in cultures of *E. coli* harboring the pMOSElox plasmid without cDNA-insert (data not shown).

The catalytic properties of recombinant 20-oxidase were exactly the same as previously found for native 20-oxidase from

pumpkin endosperm (6) and recombinant 20-oxidase from pumpkin cotyledons (7). These three enzymes catalyze the successive oxidation of the C-20 methyl group of [ $^{14}\text{C}$ ]GA<sub>12</sub> or [ $^{14}\text{C}$ ]GA<sub>53</sub> to a carboxylic acid group via an alcohol and an aldehyde (data not shown). They also convert [ $^{14}\text{C}$ ]GA<sub>12</sub> at higher rates than [ $^{14}\text{C}$ ]GA<sub>53</sub> and produce C<sub>19</sub> GAs in low yields with both precursors (data not shown).

**Sequence Analysis of cDNA Clones.** DNA sequencing of inserts of clones D4 (GA 7-oxidase) and E5 (GA 20-oxidase) identified ORFs of 314 and 386 amino acids, respectively (Fig. 3). The length of the clone D4 was confirmed by 5' RACE. The coding region of clone E5 is different at seven positions (176, 502, 545, 672, 674, 931, and 1098) from the 20-oxidase previously cloned from pumpkin cotyledons (ref. 7, not illustrated in Fig. 3), resulting in changes of the amino acids 174 (Pro to Ser), 216 (Phe to Tyr), 217 (Glu to Lys), and 358 (Asp to Ala). However, the noncoding regions of the two 20-oxidases differ considerably. The calculated molecular mass was 35,712 Da for clone D4, and pI was calculated to be 5.7. Both clones, D4 and E5, show highly conserved regions typical for 2-oxoglutarate-dependent dioxygenases, including three metal-binding residues (15, 23), His-244, Asp-246, and His-299 (numbering follows GA 20-oxidase). However, identity of the two enzymes at the amino acid sequence level is only about 25%.

## DISCUSSION

Evidence has been presented for isolation of the first GA 7-oxidase cDNA (clone D4) and an additional GA 20-oxidase cDNA (clone E5) by a novel cloning strategy. The results show that eukaryotic (plant) genes can be cloned by assaying pools of *E. coli* harboring a cDNA library derived from the eukaryotic species for heterologous expression of enzyme activity. This strategy uses standard techniques, but they do not appear to have been used previously in the combination described here. The main advantage of the strategy is that no sequence information or antibody preparation from purified protein are required for screening. An additional advantage is that the identity of the clone is determined *ab initio*, "false positive" clones being discriminated by the specificity of the enzyme assay. The technique uses a commercially available high-expression vector in the most common bacterial host. Although fungal, animal, and plant cells are often preferred (24) *E. coli* is for many purposes the best host and often very successful for heterologous expression of eukaryotic genes (25,

Table 2. Dependency of metabolism of [ $^{14}\text{C}$ ]GA substrates on pH and lysate volumes from *E. coli* transformed with pMOSElox clone D4

pH	Substrate	Lysate, $\mu\text{l}$	Products, % of recovered radioactivity							
			GA <sub>14</sub> -ald.	GA <sub>14</sub>	GA <sub>12</sub> -ald.	GA <sub>12</sub>	Compound W	Compound X	Compound Y	Compound Z
7.4	GA <sub>12</sub> -ald.	700*			0 <sup>†</sup>	82	9	9	0	0
	GA <sub>14</sub> -ald.	700*	3 <sup>†</sup>	97						
5.4	GA <sub>12</sub> -ald.	0.27			31 <sup>†</sup>	69	0	0	0	0
		1.09			15 <sup>†</sup>	83	0	0	2	0
		4.38			2 <sup>†</sup>	90	2	1	5	0
		17.5			0 <sup>†</sup>	73	5	13	7	2
		35			0 <sup>†</sup>	52	8	32	8	0
		70			0 <sup>†</sup>	27	9	58	3	3
		700*			0 <sup>†</sup>	42	7	51	0	0
		700*				31 <sup>†</sup>	7	52	10	0
	GA <sub>12</sub>	700*								
	Compound X	70					95 <sup>†</sup>		5	
	GA <sub>14</sub> -ald.	700*	0 <sup>†</sup>	100						

Lysates were prepared as described in *Material and Methods* using standard lysis buffer or a modified lysis buffer containing 200 mM BisTris-HCl (pH 5.4) instead of 200 mM Tris-HCl (pH 7.4). All products, except for compound Y and compound Z, were analyzed by GC-MS as their methyl ester trimethylsilyl ether derivatives and identified, where possible, by comparison of their mass spectra with those of published spectra (22). *m/z* (% relative abundance) for GA<sub>12</sub>: 368(2.6), 360(0.92), 336(17), 328(8), 308(100), 300(58), 246(43), 240(21); for GA<sub>14</sub>: 456(4.2), 448(1.2), 424(30), 416(11), 396(25), 388(7), 306(96), 298(65); for compound W (putative 12-hydroxy-GA<sub>12</sub>): 456(5.4), 448(4.5), 424(56), 416(27), 396(44), 388(25), 334(46), 326(27), 306(100), 239(72); for compound X: 456(5.5), 448(2.8), 424(37), 416(23), 334(21), 326(9), 291(22), 283(15), 245(100), 239(48).

\*Total volume was 10 times that of the standard assay.

<sup>†</sup>Unmetabolized substrate.

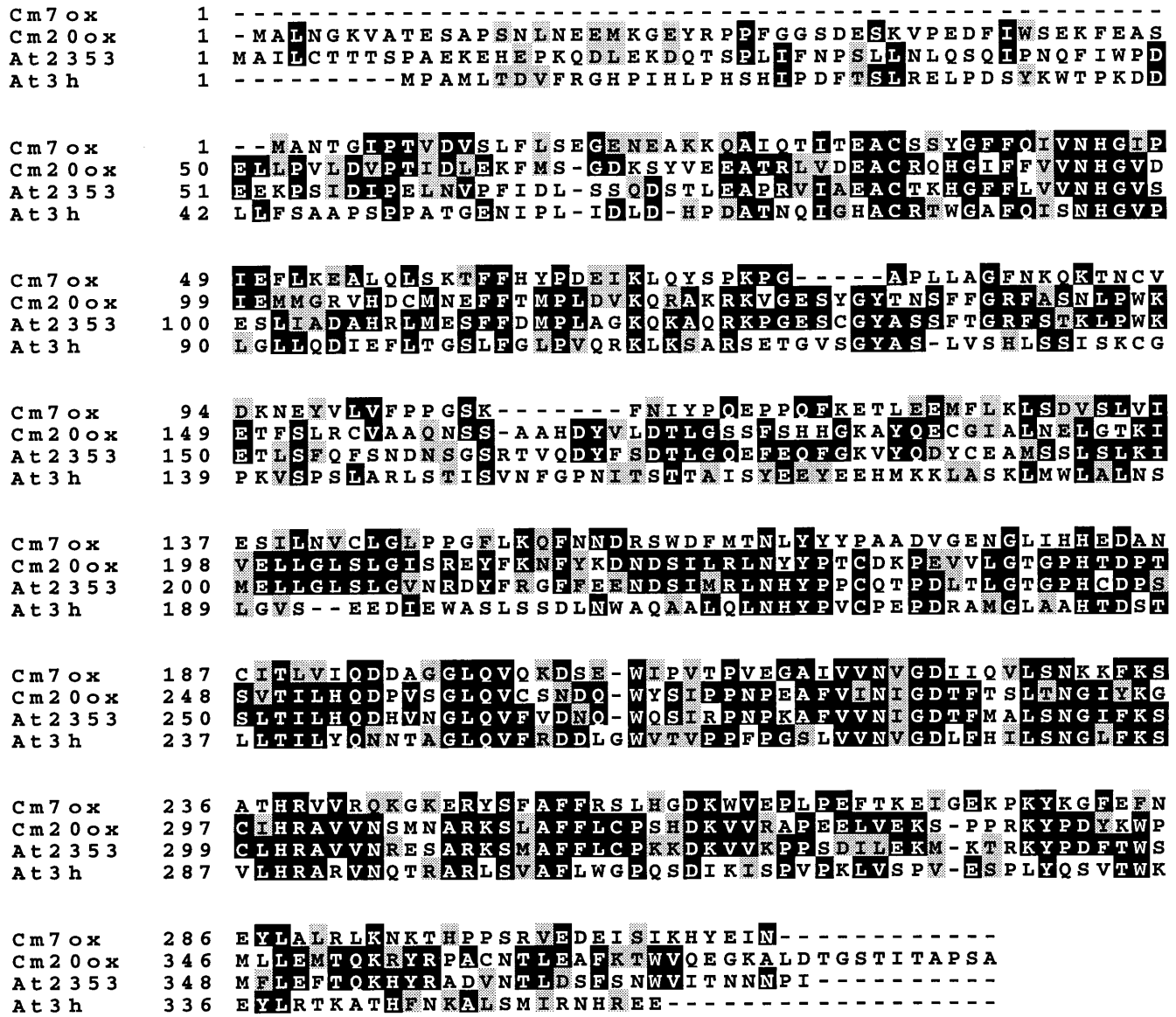


FIG. 3. Alignment of derived GA 7-oxidase (Cm7ox) and 20-oxidase (Cm20ox) amino acid sequences from pumpkin endosperm with GA 20-oxidase (At2353) (9) and putative 3β-hydroxylase sequence (At3h) (16) from Arabidopsis. Identical residues are boxed in black; similar residues are shaded in gray.

26). There are many reports on the expression of recombinant enzymes retaining catalytic activity as fusion proteins (7–15, 26–29). Therefore, the method described here should be applicable for cloning a wide range of enzymes, provided that these are not involved in the metabolism of or are toxic to *E. coli* itself. Screening is facilitated by using a pooling procedure originally developed for PCR-based screening (19).

The efficiency of this cloning approach depends on the availability of sensitive functional assays. The GA 7-oxidase clone was detected in a primary pool of 40,000 clones (Fig. 2c and Table 2) and, with multiple primary pools, up to 2,400,000 clones can be screened per person per day. Such a throughput is comparable to conventional cloning techniques using antibody or DNA probes. Because a pool of *E. coli* BL21 (DE3) transformed with the pumpkin endosperm cDNA-library expressed fusion protein with GA 3β-hydroxylase activity, in addition to 7- and 20-oxidase activities, the cloning of the 3β-hydroxylase cDNA also would have been possible by this approach. However, this possibility was not pursued, because the cDNA for the 3β-hydroxylase already had been cloned by

conventional immunoscreening (T.L. and S. Robatzek, unpublished data).

The construction of the cDNA library was optimized by choosing mRNA pools expressing high GA-dioxygenase activity in *in vitro* translation products (Table 1). The results indicate that transcript levels for the GA 7-oxidase, the GA 20-oxidase, and the GA 3β-hydroxylase were highest in the endosperm of seeds without developed cotyledons. Vegetative tissues from pumpkin contain only 1/200th to 1/6,000th of these transcript levels (A. Frisse and T.L., unpublished results). A similar seed specific expression was observed with GA 20-oxidase from *Marah macrocarpus*, a species related to pumpkin (13).

GA 7-oxidase activity in pumpkin seeds is catalyzed by two classes of enzymes, microsomal NADPH-dependent P450-monoxygenases and soluble 2-oxoglutarate-dependent dioxygenases (5, 30). The results confirm that the recombinant 7-oxidase described here belongs to the latter class of enzymes because its derived amino acid sequence contains elements highly conserved in dioxygenases (Fig. 3). Its theoretical

molecular mass is similar to the native 7-oxidase of the dioxygenase type (34.5 kDa) as determined by gel-filtration HPLC (5). The length of the cDNA clone was confirmed by 5' RACE, indicating that a full-length clone has been isolated. Thus this enzyme is about 10 kDa smaller than other GA dioxygenases (20). The major difference occurs at its N-terminal end where the 7-oxidase sequence is about 50 amino acids shorter than sequences of other known GA dioxygenases (Fig. 3). This may reflect the limited importance of these amino acids for 2-oxoglutarate dependent dioxygenase function *per se*. However, because other 20-oxidases have conserved amino acids within this region it may encode for some specific functions of individual GA-dioxygenases.

The GA 7-oxidase fused to the first 260 amino acids of the T7 gene 10 product catalyzes the oxidation of GA<sub>12</sub>-aldehyde to GA<sub>12</sub> and GA<sub>14</sub>-aldehyde to GA<sub>14</sub>. Less efficiently, it converts GA<sub>12</sub> to (putative) 12-hydroxy-GA<sub>12</sub> (compound W) and three unidentified compounds, X, Y, and Z. These same four products also were obtained in low yield from incubation of GA<sub>12</sub>-aldehyde and the native purified 7-oxidase from endosperm of pumpkin (data not shown). Mass spectra indicate that compound X represents a monohydroxylated GA<sub>12</sub>. It accumulates at high lysate concentrations and is a precursor for compound Z, which might indicate the existence of up to now unknown GA-biosynthetic pathways in the pumpkin endosperm. Thus the 7-oxidase in pumpkin endosperm appears to be multifunctional, catalyzing oxidation of an aldehydic function and the hydroxylation of GA<sub>12</sub>. Further characterization of this multifunctionality is under investigation.

A GA 20-oxidase cDNA previously has been cloned from pumpkin cotyledons (7), and the cloning of another 20-oxidase from pumpkin endosperm was reported, without details of the nucleotide sequence, but stating that it differed from the first by one amino acid (8). The coding regions of the cDNA sequences of the 20-oxidases from the endosperm, described here, and from cotyledons (7), are highly conserved. However, they differ by seven nucleotides, resulting in four changes of the derived amino acid sequence. These changes lay all within the regions of greatest variations of dioxygenase sequences as proposed by Roach *et al.* (23) and do not appear to affect the catalytic properties of the recombinant 20-oxidase. In contrast, the noncoding regions of the two sequences are very different. This could indicate the existence of a multiple gene family or, because pumpkin is believed to be an ancient tetraploid genus (31), of homeologous genes from different ancestral genomes of pumpkin.

I thank Prof. H. J. Fritz and Dr. B. Fahrman (University of Göttingen, Germany) for help with automatic sequencing, Prof. J. MacMillan, Dr. P. Hedden, and Dr. A. Phillips (Long Ashton Research Station, U.K.) for their comments on the manuscript, and the Deutsche Forschungsgemeinschaft for support. This work is dedicated to Prof. J. E. Graebe.

1. Crozier, A., ed. (1983) *The Biochemistry and Physiology of Gibberellins* (Praeger, New York), Vol. 2.
2. Graebe, J. E. (1987) *Annu. Rev. Plant Physiol.* **38**, 419–465.

3. Lange, T., Hedden, P. & Graebe, J. E. (1993) *Planta* **189**, 340–349.
4. Lange, T., Hedden, P. & Graebe, J. E. (1993) *Planta* **189**, 350–358.
5. Lange, T., Schweimer, A., Ward, D. A., Hedden, P. & Graebe, J. E. (1994) *Planta* **195**, 98–107.
6. Lange, T. (1994) *Planta* **195**, 108–115.
7. Lange, T., Hedden, P. & Graebe, J. E. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 8552–8556.
8. Xu, Y.-L., Li, L., Wu, K., Peeters, A. J. M., Gage, D. A. & Zeevaert, J. A. D. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 6640–6644.
9. Phillips, A. L., Ward, D. A., Uknes, S., Appleford, N., Lange, T., Huttly, A. K., Gaskin, P., Graebe, J. E. & Hedden, P. (1995) *Plant Physiol.* **108**, 1049–1057.
10. Wu, K., Li, L., Gage, D. A. & Zeevaert, J. A. D. (1996) *Plant Physiol.* **110**, 547–554.
11. Martin, D. N., Proebsting, W. M., Parks, T. D., Dougherty, W. G., Lange, T., Lewis, M. J., Gaskin, P. & Hedden, P. (1996) *Planta* **200**, 159–166.
12. Toyomasu, T., Kawaide, H., Sekimoto, C., von Numers, C., Phillips, A. L., Hedden, P. & Kamiya, Y. (1997) *Physiol. Plant.* **99**, 111–118.
13. MacMillan, J., Ward, D. A., Phillips, A. L., Sánchez-Beltrán, M. J., Gaskin, P., Lange, T. & Hedden, P. (1997) *Plant Physiol.* **111**, 1369–1377.
14. Hedden, P. & Kamiya, Y. (1997) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**, 431–460.
15. Prescott, A. G. & John, P. (1996) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**, 245–271.
16. Chiang, H.-H., Hwang, I. & Goodman, H. M. (1995) *Plant Cell* **7**, 195–201.
17. Graebe, J. E. (1972) in *Plant Growth Substances 1970*, ed. Carr, D. J. (Springer, Heidelberg), pp. 151–157.
18. Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557–580.
19. Israel, D. I. (1993) *Nucleic Acids Res.* **21**, 2627–2631.
20. Lange, T. & Graebe, J. E. (1993) in *Methods in Plant Biochemistry*, ed. Lea, P. J. (Academic, London), Vol. 9, pp. 403–430.
21. Homann, V., Mende, K., Arntz, C., Hardi, V., Macino, G., Morelli, G., Böse, G. & Tudzynski, B. (1996) *Curr. Genet.* **30**, 232–239.
22. Gaskin, P. & MacMillan, J. (1992) *GC-MS of the Gibberellins and Related Compounds* (Cantock's Enterprises, Bristol, U.K.).
23. Roach, P. L., Clifton, I. J., Fulop, V., Harlos, K., Barton, G. J., Hajdu, J., Andersson, I., Schofield, C. J. & Baldwin, J. E. (1995) *Nature (London)* **375**, 700–704.
24. Frommer, W. B. & Ninnemann, O. (1995) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **46**, 419–444.
25. Gold, L. (1990) *Methods Enzymol.* **185**, 11–14.
26. Makrides, S. C. (1996) *Microbiol. Rev.* **60**, 512–538.
27. Kaufman, D. L., McGinnis, J. F., Krieger, N. R. & Tobin, A. J. (1986) *Science* **232**, 1138–1140.
28. Sato, T. & Theologis, A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6621–6625.
29. Yamaguchi, S., Saito, T., Abe, H., Yamane, H., Murofushi, N. & Kamiya, Y. (1996) *Plant J.* **10**, 203–213.
30. Graebe, J. E. & Hedden, P. (1974) in *Biochemistry and Chemistry of Plant Growth Regulators*, eds. Schreiber, K., Schütte, H. R. & Sembdner, G. (Acad. Sci. G. D. R., Inst. Plant Biochem., Halle, Germany), pp. 1–16.
31. Robinson, R. W. & Decker-Walters, D. S. (1997) in *Crop Production Science in Horticulture*, eds. Atherton, J. & Rees, A. (CAB International, Oxon, U.K.), Vol. 6.