Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 3821–3825, May 1990
Biochemistry

Molecular cloning and characterization of *GPA1*, a G protein α subunit gene from *Arabidopsis thaliana*

(polymerase chain reaction/GTP-binding protein/α subunit/restriction fragment length polymorphism mapping)

HONG MA, MARTIN F. YANOFSKY, AND ELLIOT M. MEYEROWITZ*

Division of Biology, 156-29, California Institute of Technology, Pasadena, CA 91125

Communicated by Melvin I. Simon, March 8, 1990

ABSTRACT **We have isolated a gene coding for a G protein α subunit from the flowering plant *Arabidopsis thaliana*. This gene, named *GPA1*, was isolated by using a DNA probe generated by polymerase chain reaction based on protein sequences from mammalian and yeast G protein α subunits. The sequences of genomic and cDNA clones indicate that *GPA1* has 14 exons, and the deduced amino acid sequence shows that the *GPA1* gene product (GPα1) has 383 amino acid residues (44,582 Da). The GPα1 protein exhibits similarity to all known G protein α subunits—36% of its amino acids are identical and 73% are similar (identical and conservative changes) to mammalian inhibitory guanine nucleotide-binding regulatory factor α subunits and transducins. Furthermore, the GPα1 protein has all of the consensus regions for a GTP-binding protein. The *GPA1*-encoded mRNA of 1.55 kilobases is most abundant in vegetative plant tissues, as determined by RNA blot analysis. Restriction fragment length polymorphism mapping experiments show that *GPA1* is ≈1.2 centimorgans from the visible marker *er* on chromosome 2.**

Plant cells respond to a large number of external and internal stimuli, such as light, gravity, microbes, and hormones. Although much information has been obtained through physiological and biochemical studies of these responses (e.g., refs. 1–4), little is known about the molecular mechanisms of plant signal-transduction pathways. One approach to understanding signal-transduction pathways in plants is to study mutants that are defective in signal responses, and a large number of these types of mutants have been isolated in *Arabidopsis* (4–7). Analyses using these mutants have been informative about the physiology and genetics, but not yet about the mechanisms, of some signal-transduction pathways in *Arabidopsis*.

An alternative approach to the understanding of plant signaling processes is to study homologues of proteins known to play important roles in signal transduction in animals and simple eukaryotes, such as G proteins. G proteins are members of a specific family of guanine nucleotide-binding regulatory proteins that participate in a variety of signaling processes of eukaryotic organisms, from yeasts to humans (for reviews, see refs. 8–10). These heterotrimeric proteins (α, β, and γ subunits) are associated with the cytoplasmic side of cell membranes and transmit signals from transmembrane receptors to effector proteins, which, in turn, produce, often through a cascade of reactions, changes in cellular metabolism. G protein α subunits bind guanine nucleotides and have a GTPase activity. The best characterized α subunits are from mammals: stimulatory and inhibitory guanine nucleotide-binding regulatory proteins (G_s and G_i) are involved in the stimulation or inhibition of adenylate cyclase activity, respectively; transducin (T_r) is responsible for signal transduction in the rod photoreceptor from rhodopsin to cGMP phosphodiesterase. More recently discovered α subunits include G_o, of unknown function (11), found in mammalian brain and heart, and transducin from the bovine cone photoreceptor (12). Furthermore, Strathmann *et al.* (13) demonstrated by isolating sequences of five distinctive α subunits that, in mammals, G protein α subunits are members of a diverse family with several subfamilies. In yeast, a G protein has been shown to participate in the signaling of the pheromone response, and genes for all three subunits have been isolated (14–17). G protein α subunit genes have also been isolated from *Drosophila melanogaster* (18) and *Dictyostelium discoideum* (19). Although the degrees of amino acid sequence similarity of G protein α subunits between mammals and yeast or between mammals and slime mold are moderate (40–50%), they all have several highly conserved consensus regions for guanine nucleotide binding and hydrolysis (for review, see ref. 9).

Although no G protein genes have been isolated from plants previously, there has been evidence suggesting the existence of guanine nucleotide-binding regulatory proteins in plants. By using antisera raised against a peptide conserved between mammalian G protein α subunits, other workers have identified cross-reacting proteins from the cell membranes of *Cucurbita pepo* L (20), *Arabidopsis thaliana* (21), and *Commelina communis* (21), and partially purified those from *Arabidopsis* (21). To identify and study heterotrimeric G proteins in plants, we have taken a molecular approach of cloning genes encoding G protein subunits. We chose as our experimental organism *A. thaliana*, a small flowering plant that has become a model system for plant genetics and molecular biology (22, 23). Here we report the isolation and characterization of a gene encoding a G protein α subunit from *A. thaliana*.†

MATERIALS AND METHODS

Plasmid and Plant Ecotypes. The *A. thaliana* plants used here are in either the Columbia [for polymerase chain reaction (PCR)] or Landsberg (*er*) (for cosmid and cDNA libraries) ecotypes. The plasmid pCIT1828 was constructed by inserting the longest *GPA1* cDNA fragment from the λgt10 clone (see below) into the *Eco*RI site of the Promega vector pGEM7Zf(+).

PCR. PCR was performed as described (24) by using a 53°C hybridizing temperature, plant genomic DNA, and degenerate oligonucleotides similar to those described by Strathmann *et al.* (13): (*i*) 5′-CGGATCCAA(AG)TGGAT(ACT)CA(CT)-

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.

Abbreviations: PCR, polymerase chain reaction; RFLP, restriction fragment length polymorphism; ORF, open reading frame; G_s, stimulatory G protein; G_{i1}, G_{i2}, and G_{i3}, inhibitory G proteins 1, 2, and 3, respectively.

*To whom reprint requests should be addressed.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M32887).

TG(CT)TT-3′ (same as oMP19, ref. 13), (*ii*) 5′-GGAATTC(AG)TC(AGT)AT(CT)TT(AG)TT(AGCT)AG(AG)AA-3′, and (*iii*) 5′-GGAATTC(AG)TC(AGT)AT(CT)TT(AG)TT-(CT)AA(AG)AA-3′ (similar to oMP20 and oMP21, respectively, ref. 13). The underlined sequences are restriction sites for convenient cloning of the PCR products. Oligonucleotide *i* is based on the peptide Lys-Trp-Ile-His-Cys-Phe, and oligonucleotides *ii* and *iii* are both based on the peptide Phe-Leu-Asn-Lys-Ile-Asp. Oligonucleotide *i* is in the sense, and oligonucleotides *ii* and *iii* are in the antisense orientation.

Library Screening and DNA Sequencing. A cosmid library (M.F.Y. and J. Bowman, unpublished work) containing *A. thaliana* Landsberg (*er*) nuclear DNA was screened with ^{32}P-labeled PCR fragment under stringent hybridization conditions as described (25). A λgt10 cDNA library constructed with poly(A)$^+$ RNA from young flowers of the Landsberg (*er*) ecotype (J. Bowman, personal communication) was screened using a ^{32}P-labeled 2.7-kilobase (kb) genomic *Eco*RI fragment as probe, as described (25). DNA sequencing was performed using either single- or double-stranded template procedures provided in the Sequenase kit (United States Biochemical).

RNA Blot Hybridizations. Poly(A)$^+$ RNA was isolated from different tissues of *A. thaliana* Landsberg (*er*) plants by using a described procedure (26), and similar amounts of RNAs from different tissues were electrophoretically separated on an agarose gel. The RNA was then transferred to a nylon filter (Hybond-N; Amersham) and hybridized to ^{32}P-labeled *GPA1* cDNA.

Restriction Fragment Length Polymorphism (RFLP) Mapping. RFLP mapping was done as described by Chang *et al.* (27). Two DNA fragments ≈30 kb away from *GPA1* were found to reveal DNA polymorphisms between appropriate ecotypes of *Arabidopsis*. They were used to probe genomic DNA blots of segregating progeny of three crosses that involved the visible marker *ER* (erecta; ref. 28) on chromosome 2, including one that was previously used to generate an *Arabidopsis* RFLP map (27).

RESULTS AND DISCUSSION

Isolation of *GPA1*. PCR (24) is the ideal method for isolating genes encoding G protein α subunits from plants because the level of similarity between different types of G protein α subunits is moderate and the amino acid similarity is clustered in short regions. To obtain a DNA probe for genes encoding G protein α subunits in *A. thaliana*, we carried out PCR experiments with genomic DNA and degenerate oligonucleotides. The sequences of three degenerate oligonucleotides were derived from two highly conserved hexapeptides of known mammalian and yeast G protein α subunits. Although no specific PCR product was generated by using oligonucleotides *i* and *ii*, a specific DNA product of ≈360 base pairs (bp) was generated by using oligonucleotides *i* and *iii* with genomic DNA of *A. thaliana* (Columbia ecotype) as template. This fragment was cloned, and its nucleotide sequence was determined. This fragment contains sequences potentially encoding a peptide having a high degree of similarity to mammalian and yeast G protein α subunits (data not shown). In addition, based on the comparison to known G protein α subunit sequences, the fragment contains two small introns of A + T rich sequences with appropriate intron donor and acceptor sites. Using this PCR product as a probe, we isolated several genomic cosmid clones and several overlapping cDNA clones. Fig. 1*A* shows a restriction map of the genomic region of this gene, named *GPA1* for G protein α subunit 1.

Sequence of *GPA1*. The DNA sequences of two *GPA1* cDNAs, including the longest one, cDNA1828, and of the entire *GPA1* genomic region were determined. By comparing the genomic and cDNA sequences, we deduce that the *GPA1* gene has 14 exons, ranging from 38 to 348 bases, and 13 introns, from 77 to 489 bases (Fig. 1*B*). The nucleotide sequences of cDNA1828 and of 5′ and 3′ genomic sequences are shown in Fig. 2. A large open reading frame (ORF), which begins at the translation initiation ATG codon and ends at a TGA termination codon, is 1149 nucleotides long, and the predicted protein has 383 amino acid residues (44,582 Da, Fig. 2). The protein encoded by this ORF, which we name GPα1, is probably the gene product of *GPA1*, based on amino acid sequence comparison to known G protein α subunits (see below). The region in pCIT1828 encoding GPα1 is identical to the same region in the second completely sequenced cDNA, as well as to the corresponding regions in the genomic sequence.

Within a 215-bp region 5′ to the GPα1 initiation ATG in the longest cDNA there are two additional ATG codons, which are followed by two short ORFs that could code for 20 (ORF1) and 7 (ORF2) amino acids (Fig. 2). The sequences flanking the ATGs of these small ORFs do not match the plant initiation consensus sequence: WMAACAATGGC (where W = A or T and M = C or A) (29), unlike the GPα1 ATG region, which does show similarity to the consensus sequence. The presence of short ORFs upstream of the protein coding region is atypical but has been seen for the yeast *GCN4* (30) and *HAP4* (31) genes, several mammalian oncogenes (32), and recently for the maize glutamine synthetase (33) and opaque endosperm (*o2*) cDNAs (34). It is not known whether these short ORFs are translated *in vivo*. At the 5′ end of the cDNA and extending into the genomic sequence are four repeats of the sequence $(CTT)_n$, where *n* is between 2 and 9, starting at −58, −18, 30, and 62 (Fig. 2). In addition, further upstream in the genomic sequence is an element of eight consecutive

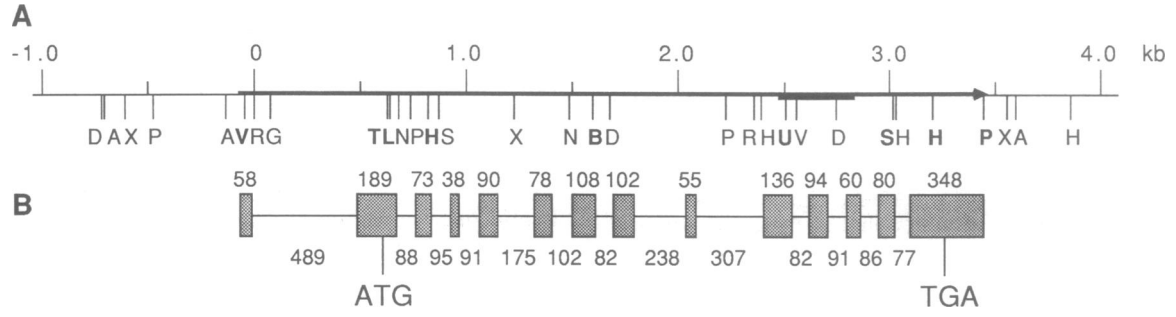

FIG. 1. (*A*) Restriction map of the *GPA1* genomic region; thick arrow indicates direction of transcription, and short bar on the genomic map denotes position of the PCR-amplified fragment. Restriction sites are derived from the DNA sequences of genomic and cDNA clones. Enzymes: A, *Sca* I; B, *Bam*HI; D, *Dra* I; G, *Bgl* II; H, *Hin*dIII; L, *Sal* I; N, *Hin*cII; P, *Hpa* I; R, *Eco*RI; S, *Sph* I; T, *Pst* I; U, *Pvu* II; V, *Eco*RV; X, *Xba* I; the sites in boldface letters are in exons. (*B*) The exon–intron structure of *GPA1*. Filled boxes are exons, including 5′- and 3′-nontranslated regions; the lines between the exons are introns. Numbers above exons and below introns are their sizes in nucleotides. Positions of the initiation ATG codon and the termination codon are indicated. *A* and *B* are scaled to each other.

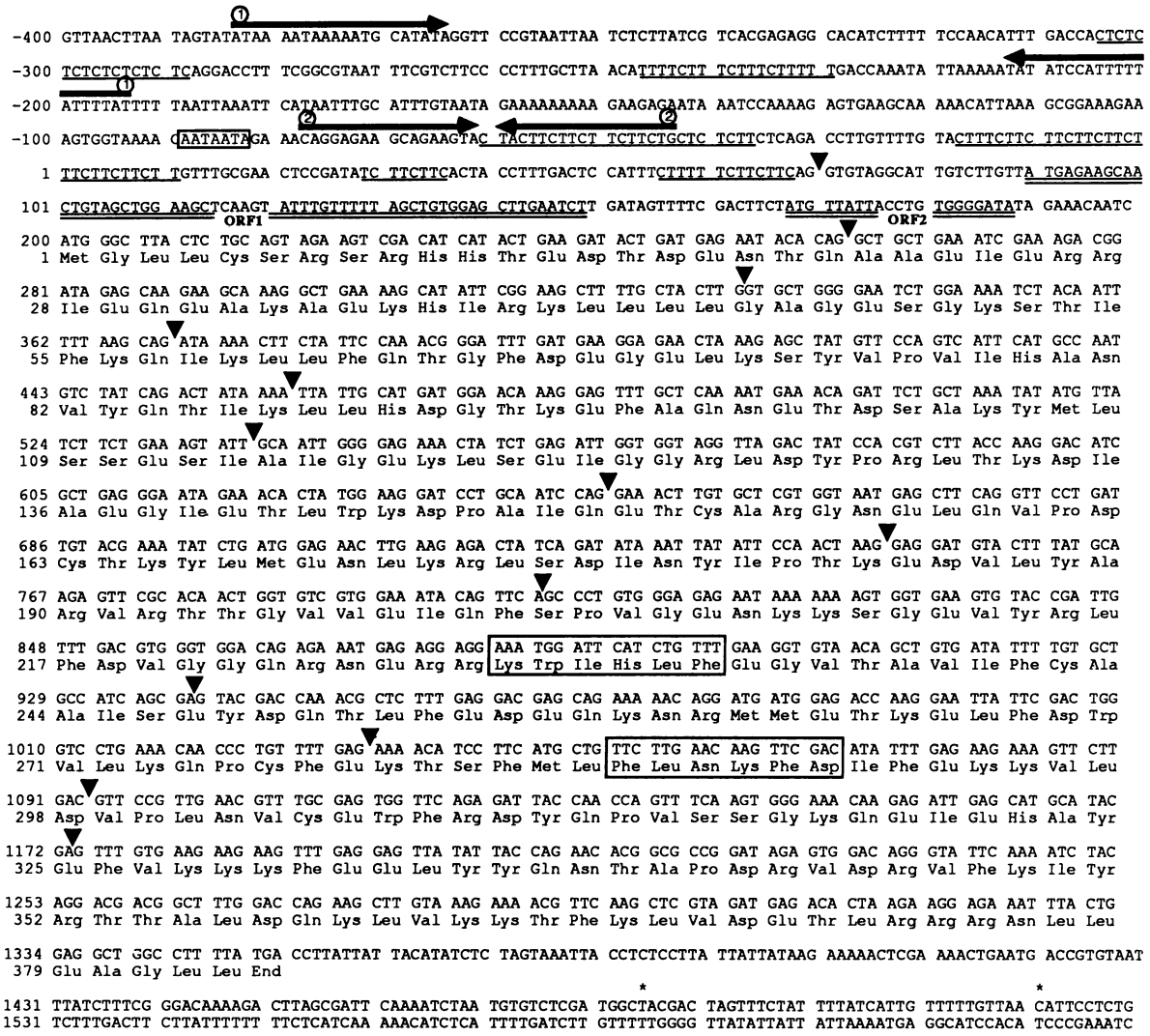

FIG. 2. Nucleotide sequence of the *GPA1* gene and the deduced amino acid sequence of GPα1 protein. Sequences of introns are not shown; ▼, positions of introns. Nucleotide 1 corresponds to the beginning of the longest cDNA. Sequences for nucleotides −400 to 11 and 1522–1830 are from genomic clones and that for 12–1521 is identical for cDNA and genomic clones. The first 11 nucleotides of the cDNA are thymidines and are believed to result from priming by the oligo(dT)-primer after first-strand synthesis. Thick arrows 1 and 2 are two pairs of inverted repeats; arrow 1 is a 19 out of 20 match, and arrow 2 is a 15 out of 17 match. The CTT and CT repeats are indicated by thin underlining. Short ORFs are double-underlined. Potential TATA sequence and two hexapeptides used to design degenerate oligonucleotides are boxed. The sequence of the PCR fragment agrees with the genomic sequence of the corresponding region. At the 3′ end of *GPA1*, two cDNAs were found to have different polyadenylylation sites, separated by 36 bp, as indicated by asterisks.

repeats of the dinucleotide CT (−305 to −289, Fig. 2). Although the significance of these repeats is unknown, similar repeated sequences have been found in the promoter regions of a number of genes (35–37). Proteins have recently been purified from *Drosophila* nuclei that bind to these elements (37, 38), and one of them was shown to stimulate transcription *in vitro* (38).

GPα1 Shares Strong Sequence Similarity with Known G Protein α Subunits. Amino acid sequence alignment reveals strong similarities of GPα1 to other G protein α subunits (Fig. 3). GPα1 has a very high degree of similarity to all members of the same family in several conserved regions; it has all of the known regions of the Holliday consensus sequences (42) for guanine nucleotide-binding (Fig. 3). Among the previously characterized G protein α subunits, rat G_{i1-3} (9) and bovine rod transducin (39) are most similar to GPα1. These proteins have 36% amino acid residues identical to GPα1, and allowing for conservative substitutions, they share ≈73% similar residues with GPα1. The genomic *GPA1* sequence was compared with those of genes encoding the human G_s (12 introns; ref. 35), rat G_{i2} and G_{i3} (8 introns; ref. 43), and *Drosophila* DGα1 (4 introns; ref. 18) proteins. Most intron positions are not conserved; however, the third *GPA1* intron was found to share the same position with the first introns of the human G_s, rat G_{i2} and G_{i3}, and *Drosophila* DGα1-encoding genes, as well as with the first intron of the rat $G_{i1}\alpha$ gene (partial sequence; ref. 43) and G_o α subunit gene (44). In addition, the fifth *GPA1* intron is at the same position as the fourth G_S intron (35), and the sixth *GPA1* intron shares the same position with the third intron of G_{i2}, G_{i3}, and G_o (35).

***GPA1* Is Differentially Expressed.** To identify the *GPA1*-encoding RNA and as a start toward determining the pattern of *GPA1* expression, a blot with RNA from several tissues was hybridized with a cDNA probe. The result (Fig. 4) indicates that there is a single mRNA for *GPA1*. Because equal amounts of poly(A)$^+$ RNAs were loaded in these lanes, the result suggests that the *GPA1* mRNA is most abundant in vegetative tissues, including leaves and/or roots, less in floral stems, and least in floral buds and floral meristem. The same blot was probed with the cDNA of a florally expressed gene (*Ag*; G. Drews, personal communication); the result shows that the floral RNA was present and not degraded. The fact

```
          *                                         ┌─A──┐
At GPA1   MGLLCS-RS-RHHTEDTDENTQAAE-IERRIEQEAKAEKHIRKLLLLGAGESGKSTIFKQIKLLFQTGFD---------------   67
Bov Tr    MGAGASA-------EEK----HSRE-LEKKLKEDAEKDARTVKLLLLGAGESGKSTIVKQMKIIHQDGYS---------------   58
Rat Go    MGCTLSA-------EERAALERSKA-IEKNLKEDGISAAKDVKLLLLGAGESGKSTIVKQMKIIHEDGFS---------------   62
Rat Gil   MGCTLSA-------EDKAAVERSKM-IDRNLREDGEKAAREVKLLLLGAGESGKSTIVKQMKIIHEAGYS---------------   62
Rat Gs    MGCLGNSKTEDQRNEEKAQREANKK-IEKQLQKDKQVYRATHRLLLLGAGESGKSTIVKQMRILHVNGFNGEGGEEDPQAARSNS   84
ystSCG1   MGCTVSTQTIGDESDPFLQNKRANDVIEQSLQLEKQRDKNEIKLLLLGAGESGKSTVLKQLKLLHQGGFS---------------   70
Dm DG1    MGCAVST------ARDKEAIERSKN-IDRALRAEGERAASEVKLLLLGAGESGKSTIVKQMKIINDTGYS---------------   63
Dd G2     MGICASS------MEGEKTNTDINLSIEKE-RKKK---NNEVKLLLLGAGESGKSTISKQMKIIHQSGYS---------------   60
                                                    GXXXXGK

At GPA1   EGELKSYVPVIHANVYQTIKLLHDGTKEFAQNETDSAKYMLSSESIAIGEKLSEIGGRLDYPRL-TKDIAEGIETLW-KDPAIQE 150
Bov Tr    LEECLEFIAIIYGNTLQSILAIVRAMTTLNIQYGDSAR-------QDDARKLMHMADTIEEGTM-PKEMSDIIQRLW-KDSGIQA 134
Rat Go    GEDVKQYKPVVYSNTIQSLAAIVRAMDTLGVEYGDKER-------KADSKMVCDVVSRMEDTEPFSAELLSAMMRLW-GDSGIQE 139
Rat Gil   EEECKQYKAVVYSNTIQSIIAIIRAMGRLKIDFGDAAR-------ADDARQLFVLAGAAEEGFM-TAELAGVIKRLW-KDSGVQA 138
Rat Gs    DGEKATKVQDIKNNLKEAIETIVAAMSNLVPPVELANP-------ENQFRVDYILSVMNVPNFDFPPEFYEHAKALW-EDEGVRA 161
ystSCG1   HQERLQYAQVIRADAIQSMKILIIQARKLGIQLDCDDP-------INN-KDLFACKRI-(117)-HEDIAKAIRGLWNNDKGIKQ 257
Dm DGA1   QEECEEYRRVVFSNTVQSLMVIIRAMGRLKIEFADPSR-------TDIARQFFTHASAADEGIL-LPEIVLLMKKLW-ADGGVQQ 139
Dd G2     NEERKEFKPIITRNILDNMRVLLDGMGRLGMPIDPSNSDAAV---MIKELTSLQASIVTDCWGELNEDQGKKIKALW-TDPGVKQ 141

                                           #                            ┌C┐
At GPA1   TCARGNELQVPDCTKYLMENLKRLSDINYIPTF[illegible]VLYARVRTTGVVEIQFSPVGENKKSGEVYRLFDVGGQRNERRKWIHLFEG 235
Bov Tr    CFDRASEYQLNDSAGYYLSDLERLVTPGYVPTEQDVLRSRVKTTGIIETQFS------FKDLNFRMFDVGGQRSERKKWIHCFEG 213
Rat Go    CFNRSREYQLNDSAKYYLDSLDRIGAADYQPTEQDILRTRVKTTGIVETHFT------FKNLHFRLFDVGGQRSERKKWIHCFED 218
Rat Gil   CFNRSREYQLNDSAAYYLNDLDRIAQPNYIPTQQDVLRTRVKTTGIVETHFT------FKDLHFKMFDVGGQRSERKKWIHCFEG 217
Rat Gs    CYERSNEYQLIDCAQYFLDKIDVIKQADYVPSDQDLLRCRVLTSGIFETKFQ------VDKVNFHMFDVGGQRDERRKWIQCFND 240
ystSCG1   CFARSNEFQLEGSAAYYFDNIEKFASPNYVCTDEDILKGRIKTTGITETEFN------IGSSKFKVLDAGGQRSERKKWIHCFEG 336
Dm DGA1   TFARSREYQLNDSAGYYLNSLDRIAQPNYIPTQQDVLRTRVKTTGIIETNFS------CKQLNFKLFDVGGQRSERKKWIHCFEG 218
Dd G2     AMRRANEFSTLPDSAPYFFDSIDRMTSPVYIPDQDILHTRVMTRGVNETNFE------IGKIKFRLVDVGGQRSERKKWLSCFDD 220
                                                                          DXXG

                 E                                      ┌─G─┐
At GPA1   VTAVIFCAAISEYDQSVFEDEQKNRMMETKELFDWVLKQPCFEKTSFMLFLNKFDIFEKKVLDVPLNVCEWFRDYQPVSSGKQEI 320
Bov Tr    VTCIIFIAALSAYDMVLVEDDEVNRMHESLHLFNSICNHRYFATTSIVLFLNKKDVFSEKIKKAHLSIC--FPDY----NGPNTY 292
Rat Go    VTAIIFCVALSGYDQVLHEDETTNRMHESLMLFDSICNNKFFIDTSIILFLNKKDLFGEKIKKSPLTIC--FPEY----PGSNTY 297
Rat Gil   VTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTDTSIILFLNKKDLFEEKIKKSPLTIC--YPEY----AGSNTY 296
Rat Gs    VTAIIFVVASSSYNMVIREDNQTNRLQEALNLFKSIWNNRWLRTISVILFLNKQDLLAEKVLAGKSKIEDYFPEF----ARYTTP 321
ystSCG1   ITAVLFVLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFILFLNKIDLFEEKVKSMPIRK--YFPDY----QGRVGD 415
Dm DGA1   VTAIIFCVALSGYDLVLAEDEENNRMIESLKLFDSICNSKWFVETSIILFLNKKDLFEEKIKRSPLTIC--FPEY----TGTNTF 297
Dd G2     VTAVVFCVALSEYDLLLYEDNSTNRMLESLRVFSDVCNS-WFVNTPIILFLNKSDLFREKIKHVDLSET--FPEY----KGGRDY 299
                                                        NKXD

                                              ┌─I─┐
At GPA1   EHAYE-------------FVKKKFEELYYQNTAPDRVDR-VFKIYRTTALDQKLVKKTFKLVDETLRRRNLLEAGLL    383
Bov Tr    EDAGN-------------YIKVQFLEL---NMRRDV--KEIY-SHMTCATDTQNVKFVFDAVTDIIIKENLKDCGLF    350
Rat Go    EDAAA-------------YIQTQFESK---NRSPN---KEIY-CHMTCATDTNNIQVVFDAVTDIIIANNLRGCGLY    354
Rat Gil   EEAAA-------------YIQCQFEDL---NKRKDT--KEIY-THFTCATDTKNVQFVFDAVTDVIIKNNLKDCGLF    354
Rat Gs    EDATPEPGEDPRVTRAKYFIRDEFLRI---STASGDGRHYCY-PHFTCAVDTENIRRVFNDCRDIIQRMHLRQYELL    394
ystSCG1   AEAGL-----------KYFEK-IFLSL---NK-TN---KPIY-VKRTCATDTQTMKFVLSAVTDLIIQQNLKKSGIIT   473
DM DGA1   EEAAN-------------YIRMKFENL---NKRKD--QKEIY-TNLTCATDTNNVKFVFDAVTDVIIKNNLKQIGLF    355
DD G2     ERASN-------------YIKERFWQI---NKTEQ---KAIY-SHITCATDTNNIRVVFEAVKDIIFTQCVMKAGLYS   357
```

FIG. 3. Alignment of GPα1 amino acid sequence with other G protein α subunit sequences. Proteins: At GPA1, *A. thaliana* GPα1; Bov Tr, bovine rod transducin (39); rat Go (rat G_o) (11), rat Gil (rat G_{i1}) (40), and rat Gs (rat G_s) (11), ystSCG1, *Saccharomyces cerevisiae* SCG1 (14, 15); Dm DG1, *Dr. melanogaster* DGα1 (18); Dd G2, *Di. discoideum* $G_\alpha 2$ (19). The amino acid sequences were aligned using the FASTP program (41). Regions A, C, E, G, and I (9, 42) are indicated, and known consensus sequences are shown below three of the regions. Solid bars indicate the positions of amino acid sequences used to design the degenerate oligonucleotides for PCR amplification. As in all known G protein α subunits, GPα1 also has a glycine residue at position 2 (★) as a site for potential *N*-myristoylation (9) and an arginine residue at position 190 (#) as a potential site for ADP-ribosylation by cholera toxin (9). Unlike G_i, G_o, and transducins, GPα1 lacks the C-terminal cysteine, the site of ADP-ribosylation by pertussis toxin (9).

that *GPA1* mRNA is more abundant in roots and/or leaves suggests that its product, GPα1, may be involved in signal-transduction pathways in one or both of these tissues. The *GPA1* mRNA is ≈1.55 kb in length, indicating that the longest cDNA clone (cDNA1828), which has 1537 bp, is nearly full length.

GPA1 Maps on Chromosome 2 Near Two Known Genes. Although GPα1 is likely to be involved in some signal-transduction pathway in *Arabidopsis*, the specific signaling

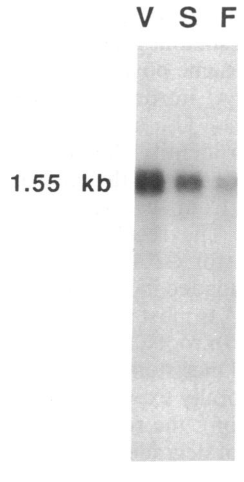

FIG. 4. Autoradiogram of an RNA blot probed with *GPA1* cDNA. Lanes: V, vegetative tissues from whole plants before bolting, including roots and leaves; S, floral stems; F, floral buds and apical meristems. Two micrograms of poly(A)$^+$ RNA was loaded in each lane.

process for which GPα1 is needed is not yet known. A large number of known mutations affect various responses to signals in *Arabidopsis*. As a first step toward determining whether or not GPα1 protein is involved in one of the known plant responses, we would like to determine whether *GPA1* is the same as one of the previously identified genes. For this purpose, we determined the genetic map position of *GPA1* using the RFLP mapping method (27). Five recombinants between *ER* and the DNA polymorphism were found among 433 meiotic products. These results localized the *GPA1*-associated polymorphisms to 1.2 centimorgans (on average, ≈150 kb) from *ER* on chromosome 2, although our data do not reveal to which side of *ER GPA1* maps. The only other known *Arabidopsis* gene in this region is *HY1*, which also maps ≈1 centimorgan from *ER* and is one of several genes (*HY1–HY5*, ref. 45; *HY6*, ref. 46) that are defined by mutations that cause the failure of homozygous mutant seedlings to respond properly to light.

Conclusion. We have isolated a gene (*GPA1*) from *A. thaliana* that encodes a G protein α subunit that shares substantial similarity with known G protein α subunits. The *GPA1* gene is differentially expressed and maps to chromosome 2 near the visible marker *ER*. Isolation of this G α subunit gene demonstrates that, like animals and simple eukaryotes, *Arabidopsis* has at least one G protein α subunit. Although high-stringency hybridizations reveal only the bands expected for the *GPA1* gene, low-stringency hybrid-

ization of an *Arabidopsis* genomic DNA blot with the *GPA1* cDNA uncovers additional bands (data not shown), suggesting the presence of other G α subunit genes in the *Arabidopsis* genome. From our knowledge of G proteins in other systems, it seems likely that there are β and γ subunits in *Arabidopsis* as well. Furthermore, it is reasonable to conclude that all plants have heterotrimeric guanine nucleotide-binding regulatory proteins. Having isolated a gene coding for a G protein α subunit, we can now begin to study the function of G proteins in plants to gain insight into the presently unknown mechanisms of plant signal transduction.

We thank M. Simon, M. Strathmann, M. Lochrie, and T. Wilkie for gifts of oligonucleotides, advice, and helpful discussions; A. van der Bliek for help with PCR experiments; J. Bowman for the cDNA library and RFLP DNA blots; G. Drews for helping with the RNA blot; S. Kempin for computer analysis of the RFLP mapping data; S. Kempin and Y. Hu for excellent technical assistance; and T. Stearns for help with the protein sequence data base search. We thank R. Bourret, L. Brockman, C. Chang, G. Drews, M. Lochrie, M. Simon, M. Strathmann, A. van der Bliek, D. Weigel, and T. Wilkie for critical reading of this manuscript. This work was supported by a grant from the National Science Foundation (DCB 8703439) and a grant from the Lucille P. Markey Charitable Trust to E.M.M. H.M. was supported by a postdoctoral fellowship from the Helen Hay Whitney Foundation, and M.F.Y. was supported by a postdoctoral fellowship in plant biology from the National Science Foundation (DCB 8608456).

1. Karssen, C. M., Brinkhorst-van der Swan, D. L. C., Breekland, A. E. & Koornneef, M. (1983) *Planta* **157**, 158–165.
2. Karssen, C. M., Groot, S. P. C. & Koornneef, M. (1987) in *Developmental Mutants of Higher Plants*, eds. Thomas, H. & Grierson, D. (Cambridge Univ. Press, Cambridge), pp. 119–133.
3. Mirza, J. I. (1987) *Plant Physiol.* **83**, 118–120.
4. Mirza, J. I., Olsen, G. M., Iversen, T.-H. & Maher, E. P. (1984) *Physiol. Plant.* **60**, 516–522.
5. Koornneef, M., Reuling, G. & Karssen, C. M. (1984) *Physiol. Plant.* **61**, 377–383.
6. Estelle, M. A. & Somerville, C. R. (1987) *Mol. Gen. Genet.* **206**, 200–206.
7. Bleecker, A. B., Estelle, M. A., Somerville, C. & Kende, H. (1988) *Science* **241**, 1086–1089.
8. Gilman, A. G. (1987) *Annu. Rev. Biochem.* **56**, 615–649.
9. Lochrie, M. A. & Simon, M. I. (1988) *Biochemistry* **27**, 4957–4965.
10. Ross, E. M. (1989) *Neuron* **3**, 141–152.
11. Itoh, H., Kozasa, T., Nagata, S., Nakamura, S., Katada, T., Ui, M., Iwai, S., Ohtsuka, E., Kawasaki, H., Suzuki, K. & Kaziro, Y. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3776–3780.
12. Lochrie, M. A., Hurley, J. & Simon, M. I. (1986) *Science* **228**, 96–99.
13. Strathmann, M., Wilkie, T. M. & Simon, M. I. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7407–7409.
14. Nakafuku, M., Itoh, H., Nakamura, S. & Kaziro, Y. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2140–2144.
15. Dietzel, C. & Kurjan, J. (1987) *Cell* **50**, 1001–1010.
16. Miyajima, I., Nakafuku, M., Nakayama, N., Brenner, C., Miyajima, A., Kaibuchi, K., Arai, K., Kaziro, Y. & Matsumoto, K. (1987) *Cell* **50**, 1011–1019.
17. Whiteway, M., Hougan, L., Dignard, D., Thomas, D. Y., Bell, L., Saari, G. C., Grant, F. J., O'Hara, P. & Mackay, V. L. (1989) *Cell* **56**, 467–477.
18. Provost, N. M., Somers, D. E. & Hurley, J. B. (1988) *J. Biol. Chem.* **263**, 12070–12076.
19. Pupillo, M., Kumagai, A., Pitt, G. S., Firtel, R. A. & Devreotes, P. N. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4892–4896.
20. Jacobs, M., Thelen, M. P., Famdale, R. W., Astle, M. C. & Rubery, P. H. (1988) *Biochem. Biophys. Res. Commun.* **155**, 1478–1484.
21. Blum, W., Hinsch, K.-D., Schultz, G. & Weiler, E. W. (1988) *Biochem. Biophys. Res. Commun.* **156**, 954–959.
22. Meyerowitz, E. M. (1987) *Annu. Rev. Genet.* **21**, 93–111.
23. Meyerowitz, E. M. (1989) *Cell* **56**, 263–269.
24. Mullis, K. B. & Faloona, F. (1987) *Methods Enzymol.* **155**, 335–350.
25. Chang, C. & Meyerowitz, E. M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1208–1212.
26. Crawford, N. M., Campbell, W. H. & Davis, R. W. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8073–8076.
27. Chang, C., Bowman, J. L., DeJohn, A. W., Lander, E. S. & Meyerowitz, E. M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6856–6860.
28. Koornneef, M. (1987) in *Genetic Maps 1987: A Compilation of Linkage and Restriction Maps of Genetically Studied Organisms*, ed. O'Brien, S. J. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 742–745.
29. Lutcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F. & Scheele, G. A. (1987) *EMBO J.* **6**, 43–48.
30. Hinnebusch, A. G. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6442–6446.
31. Forsburg, S. L. & Guarente, L. (1989) *Genes Dev.* **3**, 1166–1178.
32. Kozak, M. (1987) *Nucleic Acids Res.* **15**, 8125–8148.
33. Snustad, D. P., Hunsperger, J. P., Chereskin, B. M. & Messing, J. (1988) *Genetics* **120**, 111–124.
34. Schmidt, R. J., Burr, F. A., Aukerman, M. J. & Burr, B. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 46–50.
35. Kozasa, T., Itoh, H., Tsukamoto, T. & Kaziro, Y. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2081–2085.
36. Ishii, S., Xu, Y.-H., Stratton, R. H., Roe, B. A., Merlino, G. T. & Pastan, I. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4920–4924.
37. Gilmour, D. S., Thomas, G. H. & Elgin, S. C. R. (1989) *Science* **245**, 1487–1490.
38. Biggin, M. D. & Tjian, R. (1988) *Cell* **53**, 699–711.
39. Medynski, D. C., Sullivan, K., Smith, D., Van Dop, C., Chang, F.-H., Fung, B. K.-K., Seeburg, P. H. & Bourne, H. R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4311–4315.
40. Nukata, T., Tanabe, T., Takahashi, H., Noda, M., Haga, K., Haga, T., Ichiyama, A., Kangawa, K., Hiranaga, M., Matsuo, H. & Numa, S. (1986) *FEBS Lett.* **197**, 305–310.
41. Lipman, D. J. & Pearson, W. R. (1985) *Science* **227**, 1435–1441.
42. Holliday, K. R. (1983–1984) *J. Cyclic Nucleotide Protein Phosphorylation Res.* **9**, 435–448.
43. Itoh, H., Toyama, R., Kozasa, T., Tsukamoto, T., Matsuoka, M. & Kaziro, Y. (1988) *J. Biol. Chem.* **263**, 6656–6664.
44. Kaziro, Y., Itoh, H., Kozasa, T., Toyama, R., Tsukamoto, T., Matsuoka, M., Nakafuku, T., Obara, T., Takagi, T. & Hernandez, R. (1988) *Cold Spring Harbor Symp. Quant. Biol.* **53**, 209–220.
45. Koornneef, M., Rolff, E. & Spruit, C. J. P. (1980) *Z. Pflanzenphysiol.* **100S**, 147–160.
46. Chory, J., Peto, C. A., Ashbaugh, M., Saganich, R., Pratt, L. & Ausubel, F. (1989) *Plant Cell* **1**, 867–880.