THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 274, No. 49, Issue of December 3, pp. 34993–35004, 1999
Printed in U.S.A.
© 1999 by The American Society for Biochemistry and Molecular Biology, Inc.

Topology, Subcellular Localization, and Sequence Diversity of the Mlo Family in Plants*

(Received for publication, May 27, 1999, and in revised form, September 13, 1999)

Alessandra Devoto‡§, Pietro Piffanelli‡, IngMarie Nilsson¶, Erik Wallin¶, Ralph Panstruga‡, Gunnar von Heijne¶, and Paul Schulze-Lefert‡‖

From the ‡Sainsbury Laboratory, John Innes Centre, Colney Lane, NR4 7UH Norwich, United Kingdom and the ¶Department of Biochemistry, Stockholm University, S-106 91 Stockholm, Sweden

Barley Mlo defines the founder of a novel class of plant integral membrane proteins. Lack of the wild type protein leads to broad spectrum disease resistance against the pathogenic powdery mildew fungus and deregulated leaf cell death. Scanning *N*-glycosylation mutagenesis and Mlo-Lep fusion proteins demonstrated that Mlo is membrane-anchored by 7 transmembrane (TM) helices such that the N terminus is located extracellularly and the C terminus intracellularly. Fractionation of leaf cells and immunoblotting localized the protein to the plant plasma membrane. A genome-wide search for *Mlo* sequence-related genes in *Arabidopsis thaliana* revealed approximately 35 family members, the only abundant gene family encoding 7 TM proteins in higher plants. The sequence variability of Mlo family members within a single species, their topology and subcellular localization are reminiscent of the most abundant class of metazoan 7 TM receptors, the G-protein-coupled receptors.

Plants have developed a highly specific recognition system to mount efficient resistance responses to invading pathogens. This form of resistance is effective against single pathogen isolates and depends on plant encoded race-specific resistance (*R*) genes mediating recognition of pathogen determinants (1). The major class of *R* genes contains an N-terminal nucleotide binding site and C-terminal leucine-rich repeats of various length. A second class comprises genes containing a kinase and/or a leucine-rich repeat domain (2). In almost all cases, *R* gene-mediated resistance leads to a localized activation of host cell death thought to have a crucial role in containing pathogen ingress.

Plant pathogen resistance can also be induced by mutations in single host genes. These mutants generally exhibit resistance to single or multiple classes of pathogens. Examples are the *Arabidopsis lsd1* and *edr1* mutants, each exhibiting resistance to both virulent bacterial and fungal pathogens (3, 4). In rice, the *sl* mutant confers resistance to the rice blast fungus, *Magnaporthe grisea* (5). Similarly, the barley *mlo* mutant confers broad spectrum resistance to the powdery mildew fungus, *Erysiphe graminis* f.sp. *hordei*. Hallmarks of these mutants are cell death control perturbations in the absence of pathogens, leading to spontaneous cell death reactions in response to either developmental or abiotic cues. It seems likely that the corresponding wild type genes function as negative regulators of cell death, which define or interpret signal thresholds of plant cell death pathways.

The *mlo*-mediated resistance response in barley is dependent on at least two additional genes, *Ror1* and *Ror2* (6). Double mutants *mlo ror1* and *mlo ror2* each restore partial susceptibility to the fungus and exhibit reduced spontaneous leaf cell death. Because the double mutants affect both pathogen resistance and leaf cell death control, the genetic programs must be intertwined. Similarly, a *lsd1 eds1* double mutant in *Arabidopsis* suppresses both the spontaneous leaf cell death phenotype and pathogen resistance observed in the *lsd1* mutant.[1] Since wild type *EDS1* is known to be essential for the function of multiple *R* genes in *Arabidopsis* (7), the *lsd1 eds1* double mutant phenotype demonstrates a genetic overlap in leaf cell death control and recognition-based pathogen resistance.

A better understanding of the mechanism linking cell death control and pathogen resistance depends on the definition of biochemical functions for proteins encoded by these regulatory genes. LSD1 is a putative cytoplasmic zinc finger protein, which responds to a superoxide-dependent signal (8). Barley *Mlo* encodes the founder of a novel class of plant proteins predicted to be membrane-anchored and shown to function in a cell-autonomous fashion (9, 10).

Here we report experiments aimed at identifying possible biochemical functions of barley Mlo. We have carried out an in depth topological study of the wild type Mlo protein, as the membrane arrangement of polytopic membrane proteins is an important determinant of their biochemical function. We have used two complementary techniques, scanning *N*-glycosylation mutagenesis and Mlo-Lep fusion proteins, which enabled us to deduce a 7 TM[2] organization for the wild type protein. This structure defines a common scaffold for all sequence-related family members identified in the data base. Subcellular fractionation studies located Mlo in the plasma membrane. A genome-wide analysis of polytopic membrane proteins in *Arabidopsis* identified the Mlo family as the only abundant class with a membrane organization similar to serpentine receptors. Our data suggest that Mlo proteins may function as G-protein-coupled receptors based on common hallmarks as defined by subcellular location, membrane topology, and domain-specific sequence variability.

* This work was supported in part by grants from the Gatsby Charitable organization and the BBSRC (to P. S.-L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Supported by an EU fellowship (Training and Mobility of Researcher Program in Biotechnology).

‖ To whom correspondence should be addressed. Fax: 44-1603-250024; E-mail: schlef@bbsrc.ac.uk.

[1] J. Parker, personal communication.

[2] The abbreviations used are: TM, transmembrane; PM, plasma membrane; PAGE, polyacrylamide gel electrophoresis; Endo H, endoglycosidase H; GPCR, G-protein-coupled receptor; HS, hydrophobic segment; OST, oligosaccharyl transferase; GAS, *N*-glycosylation acceptor site; ER, endoplasmic reticulum; TMHMM, transmembrane hidden Markov model algorithm.

TABLE I
Compilation of Mlo family members

Gene designation	Species	Genome position	Length in amino acids	No. of introns	Accession no.
Mlo	*Hordeum vulgare*	Chr. 4H[a]	533	11	Z83834
Mlo2	*H. vulgare*	Chr. 4H	544	11	Y14573
OsMlo1	*O. sativa*		537	11	Z95353
AtMlo1	*A. thaliana*	Chr. IV, 15 cM	526	11	AC002330
AtMlo2	*A. thaliana*	Chr. I, 15 cM	574	13	AC007259/N37544/T88073
AtMlo4	*A. thaliana*	Chr. I, 9 cM	570	14	ATU95973
AtMlo5	*A. thaliana*	Chr. II, 66 cM	500	14	AC002332/ATU78721
AtMlo6	*A. thaliana*	Chr. I, 84.36 cM	583	13	AC005850
AtMlo7	*A. thaliana*	Chr. II, 32.5 cM	542	13	AC002329
AtMlo8	*A. thaliana*	Chr. II, 32.5 cM	593	14	AC002329
AtMlo10	*A. thaliana*	Chr. V, 128 cM	569	14	AB011474
AtMlo11	*A. thaliana*	Chr. V, 100 cM	565	12	AB017066
AtMlo15	*A. thaliana*	Chr. II, 78 cM	496	13	AC004005
AtMlo18	*A. thaliana*	Chr. II, 72.4 cM	576	13	AC004697
AtMlo20	*A. thaliana*	Chr. IV, 69 cM	478	13	AC005861
LuMlo	*L. usitatissimum*		PS[b]		AJ005341
PoMlo	*P. balsamifera*		PS		AI166945
Gmmlo	*G. max*		PS		AI440563
BrMlo	*B. rapa*		PS		AT000894
Lemlo	*L. esculentum*		PS		AI484886
CoMlo	*G. hirsutum*		PS		AI054629

[a] Chr., chromosome.
[b] PS, partial sequence.

EXPERIMENTAL PROCEDURES

Enzymes, Chemicals, and Plasmid Vectors—Unless otherwise stated, all enzymes used for *in vitro* trascription and translation reactions were obtained from Promega Biotech (Madison, WI). Plasmid pGEM3Zf (−) was obtained from Promega Biotech. The cap analog m7G(5′)-ppp(5′), the L-^{35}S-labeled amino acids methionine and cysteine and the ^{14}C-methylated size marker proteins were obtained from Amersham Pharmacia Biotech (Uppsala, Sweden). Endo H was purchased from Roche Molecular Biochemicals (Mannheim, Germany). The glycosylation acceptor peptide *N*-benzoyl-Asn-Leu-Thr-*N*-methylamide and the non-acceptor peptide *N*-benzoyl-Asn-Leu-(allo)Thr-*N*-methylamide were synthesized according to Erickson and Merrifield (11). Oligonucleotides were purchased from Genosys (United Kingdom (UK)) and Cybergene AB (Sweden).

Construction of Plasmids—All DNA manipulations were performed using *Escherichia coli* strains DH10B and JM110 (12, 13). All constructs described in this work were confirmed by DNA sequencing using an Applied Biosystem 377 automatic sequencer.

Mlo cDNA in pBluescript II KS+ (pRT2E) (9) was excised as a *Hin*dIII/*Eco*RI fragment and cloned into the vector pGEM3Zf(−) to obtain the plasmid pADN1. The 5′ end of the *Mlo* gene was modified by changing the context immediately 5′ to the initiator ATG codon to a "Kozak consensus" sequence (14). The modified *Mlo* gene was cloned in pGEM3Zf(−) to obtain plasmid pGO. The 5′ region of the gene was modified to: . . . CTCAAGCTTGCCACCATG . . . (HindIII restriction site and ATG initiator codon underlined). Primers ADUP7 (5′-ATA GAA GCT TGC CAC CAT GTC GGA CAA AAA AGG GG-3′) and ADDOWN6 (5′-CTA GAA TTC TCA TCC CTG GCT GAA GG-3′) were used to amplify the entire coding region of *Mlo* from pRT2E for cloning into plasmid pGEM3Zf(−). The plasmid pGOH, containing a 3′ truncated *Mlo*, was created by exploiting the presence of a *Pst*I site at nucleotide position 1,165 relative to ATG.

Plasmid pGO was used as template to construct all *Mlo* variants tested by *in vitro* transcription/translation. Each mutation creating an *N*-glycosylation acceptor site (GAS) was introduced by the splicing by overlap extension method (15). Plasmid pN-term was obtained by substitution of a *Hin*dIII/*Sac*I fragment in pGO; plasmids pL2A, pL2B, pL1, and pL3 were obtained by the substitution of a *Hin*dIII/*Pst*I fragment in pGO; plasmid pL5 was obtained by the substitution of a *Hin*cII/*Pst*I fragment in pGO; plasmids pC-termA and pC-termB were obtained by the substitution of a *Hin*dIII/*Eco*RI fragment in pGO. Primers and the mutations introduced to obtain the above plasmids are described in Table II.

Plasmid pGEM1 carrying a modified *lep* gene as a *Xba*I-*Sma*I fragment behind the SP6 promoter (16) was used for the construction of *Lep-Mlo* fusion constructs. In the modified *lep* gene, the nucleotide sequences coding for the wild type GAS at position 214–216 was changed to the non-acceptor sequence Gln-Gln-Thr. A novel Asn-Ser-Thr GAS was introduced at codons 90–92 (17). *Bcl*I and *Nde*I restriction sites were introduced flanking to the H2 region in codons 59 and 80, respectively. Different fragments of the barley *Mlo* gene were amplified by polymerase chain reaction from plasmid pGO using sense primers containing a *Bcl*I or a *Xba*I restriction site in combination with the antisense primer containing a *Nde*I restriction site. The sense primers used were XbaMlo (5′-ATA CTC TAG AGC CAC CAT GTC GGA C-3′) and BclMlo (5′-TGA CTG ATC AGG TGG GAG GCG CTG GAG-3′); the antisense primer used was NdeMlo (5′-TAC TCA TAT GTG CTC AGC CCG ATC TG-3′). The presence of a *Bcl*I site at nucleotide position 898 of *Mlo* was exploited to obtain plasmids pB5N6 and pXB5. To obtain plasmid pB5N6, the H2 domain of Lep was substituted by the hydrophobic segment (HS) VI of *Mlo* covering the peptide from Ile^{301} to Ser^{375}. To obtain plasmid pB1N4, the H2 domain of Lep was substituted by an *Mlo* fragment encompassing HS II–IV from Trp^{54} to Asn^{285}. The plasmid pXN4 contains the P2 domain of Lep fused downstream of Mlo HS IV encompassing Met^{1} to Asn^{285}. The plasmid pXB5 contains the H2 and the P2 domains of the Lep fused downstream of Mlo HS V encompassing Met^{1} to Ile^{301}.

In Vitro Transcription/Translation in the Presence of Reticulocyte Lysate—Supercoiled pGEM1 and pGEM3Zf (−) plasmids containing the relevant *Mlo* derivatives were purified using the Plasmid maxi kit (Qiagen, West Sussex, UK). The supercoiled plasmids were used directly for *in vitro* transcription with the exception of plasmid pGOH, which was linearized with the restriction enzyme *Pst*I. Messenger RNA was transcribed in a 50-μl reaction volume from the SP6 promoter using 1–5 μg of DNA template. The reaction mixture contained SP6 transcription buffer (40 mM Hepes-KOH, pH 7.4, 6 mM MgOAc, 2 mM spermidine), 1 mM ATP, 1 mM CTP, 1 mM UTP, 0.5 mM GTP, 1 mM m7G(5′)-ppp(5′) 20 units of RNA polymerase, 5 μg of bovine serum albumin, 5 mM dithiothreitol, and 60 units of RNasin. The reactions were carried out at 37 °C for 1 h. The transcription of the plasmid pGOH, containing the 3′ truncated *Mlo*, was carried out using the RiboMAX large scale RNA production system SP6 following manufacturer's instructions. The reactions were carried out at 30 °C for 12 h. The RNA from each transcription was analyzed by agarose gel electrophoresis to examine size and homogeneity. Messenger RNA was then purified using the RNeasy Plant mini kit (Qiagen) prior to the translation reactions.

In vitro translation of L-[^{35}S]methionine or L-[^{35}S]cysteine labeled proteins from *in vitro* synthesized mRNA was performed by modifying the protocol described by Melancon and Garoff (18). Nine μl of rabbit reticulocyte lysate were added to the translation mixture, with or without 1 μl of dog pancreas ER microsomal membranes. The translation mixture consisted of 1 μl of the *in vitro* synthesized RNA, 40 units of RNasin, 0.03 mM amino acid mixture minus methionine or minus methionine and cysteine, 10 μCi of L-[^{35}S]methionine, or a mixture of L-[^{35}S]methionine and L-[^{35}S]cysteine in final volume of 30 μl. Translations were carried out at 30 °C for 1 h. The occurrence of *N*-linked glycosylation was tested in the presence of competing (acceptor) or control (non-acceptor) tripeptides to a final concentration of 200 μM as

described previously (19). Alternatively, they were assayed by endoglycosidase H (Endo H) treatment of the translation mixture. Reactions were performed with 50 milliunits of Endo H in 1% (w/v) SDS, 50 mM sodium citrate, pH 6.0, for 6 h at 37 °C. Membrane integration of the translation products was determined by extraction with 90 μl of ice-cold 0.1 M Na_2CO_3, pH 11, followed by recovery of the stripped microsomes by centrifugation (245,000 × *g* for 10 min). For protease protection experiments, proteinase K (Sigma, Dorset, UK) or trypsin (Sigma) were added to a final concentration of 0.5 and 1 mg/ml, respectively. The samples were incubated on ice for 30 min in presence of 10 mM $CaCl_2$ to preserve membrane integrity. Proteolysis was stopped by adding phenylmethylsulfonyl fluoride (BDH, Dorset, UK) to a final concentration of 400 μg/ml and trypsin inhibitor (Sigma) to a final concentration of 20 μg/ml. Samples were kept on ice for an additional 5 min before being processed for gel electrophoresis. All samples were analyzed by using 10% or 12% SDS-PAGE. The gels were scanned in a Fujix BAS 1000 phosphorimaging plate scanner and analyzed using the MacBas software (version 2.31).

Plant Material—Plant primary leaf material was harvested from 8-day-old seedlings grown in a controlled environment with a 16-h photoperiod at 15 °C. The analyzed barley genotypes cultivar Ingrid *Mlo*, BC Ingrid *mlo-3*, and BC Ingrid *mlo*-4 have been described previously (9).

Antibody Generation and Immunoblot Analysis—The MYP multiple antigenic peptide (^{493}MYPVVVAHPVHRL^{505} in barley Mlo) was synthesized using Fmoc solid-phase peptide synthesis on an Applied Biosystems "Synergy" model 432A peptide synthesizer (Ref. 20; Applied Biosystems, Warrington, Cheshire, UK). The MYP polyclonal antibody was raised in a New Zealand White rabbit, and the serum was affinity-purified. MYP affinity column purification of the antibody was carried out using the AminoLink Plus immobilization kit (Pierce, Chester, UK). The purified MYP antibody was dialyzed against 0.5× phosphate-buffered saline (1.3 M NaCl, 0.07 M $Na_2HPO_4 \cdot 12H_2O$, 0.03 M $NaH_2PO_4 \cdot H_2O$, pH 7.3) for 16 h. A dilution of 1:1,000 was used in all immunoblot experiments. The anti-maize calreticulin antibody was kindly donated by Dr. R. Napier (21), and the 759 PM ATPase antibody was provided by Dr. R. Serrano (22).

Membrane Vesicle Fractionation—1 g of primary barley leaves were cut and ground on ice in 10 ml of ice-cold lysis buffer (20 mM Hepes-KOH, pH 7.5, 13% (w/v) sucrose, 1 mM dithiothreitol, 1 mM EDTA, and one tablet of Complete protease inhibitor mixture (Roche Molecular Biochemicals)). The cell homogenate was filtered through two layers of Miracloth (Calbiochem, Nottingham, UK) to remove unbroken cells and was centrifuged at 3,000 × *g* for 5 min at 4 °C. For differential centrifugation experiments, the supernatant was recovered and further centrifuged at 10,000, 20,000, and 50,000 × *g* for 20 min and then at 100,000 × *g* for 45 min, with the supernatant from each spin taken for use in the next centrifugation step. The resulting pellets following each centrifugation step were resuspended in cold lysis buffer. Proteins were quantified according to the method of Bradford (23) using bovine serum albumin as a standard. Protein extracts were incubated in SDS denaturing buffer (200 mM Tris-HCl, 20% (w/v) glycerol, 5 mM EDTA, 0.02% (w/v) bromphenol blue, 2% (w/v) SDS) for 10 min at 55 °C, 50 μg/lane separated by 10% SDS-PAGE, and transferred to Hybond-C-ECL (Amersham Pharmacia Biotech) nitrocellulose membranes. Blots were stained with Ponceau S (Sigma) and immersed in blocking solution (5% (w/v) milk powder in 1× Tris-buffered saline (10 mM Tris-HCl pH 7.5, 150 mM NaCl) and 0.05% (w/v) Tween 20) for 2 h at room temperature. Incubation with the MYP antibody was carried out for 4 h at 25 °C, followed by washing in 1× Tris-buffered saline for 30 min. Incubation with the anti-rabbit peroxidase secondary antibody (1:10,000) and luminography was carried out according to the manufacturer's instructions (Amersham Pharmacia Biotech).

Purification of Plasma Membrane Vesicles by Two-phase Aqueous Partitioning—The method is based on the observation that vesicles derived from different membranes are separated into different polymer phases as a result of their unique surface properties (24). Primary leaves from 8-day-old barley seedlings (60 g) were used for the plasma membrane preparations as described by Larsson (25) with minor modifications. Four cycles of two-phase partitioning were carried out to reduce ER contamination. The plasma membrane vesicles were delipidated by mixing 100 μl of membranes with 400 μl of methanol, 100 μl of chloroform, and 500 μl of distilled water. The emulsion was mixed for 15 s and centrifuged for 5 min at 100,000 × *g* (4 °C). 400 μl of methanol were added to the recovered upper phase, centrifuged for 5 min at 13,000 × *g* (4 °C). The pellet was then resuspended in SDS sample buffer (4% (w/v) SDS, 200 mM Tris-HCl, pH 8.8, 20% (w/v) glycerol, 5 mM EDTA, 0.2% (w/v) bromphenol blue, 5% (w/v) β-mercaptoethanol).

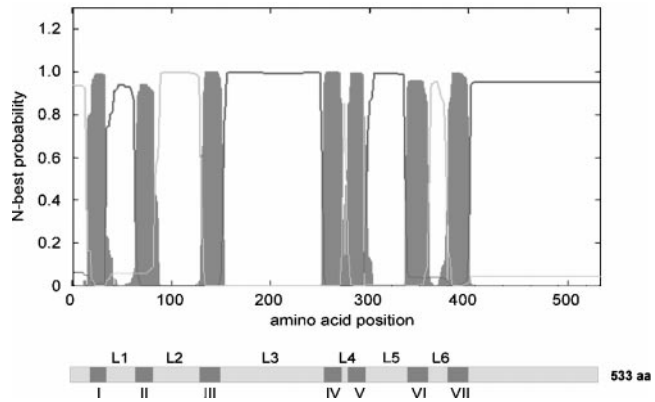

FIG. 1. **Predicted topology of barley Mlo.** The *upper part* shows a graphic representation of N-best probabilities calculated for extracellular (*light gray line*) and intracellular (*dark gray line*) Mlo domains based on the TMHMM algorithm (28). *Gray columns* depict hydrophobic segments likely forming transmembrane helices. The *lower part* of the figure graphically summarizes the TMHMM analysis for the 533-amino acid Mlo protein. *Roman numerals* indicate the position of hydrophobic segments (*I–VII*); loops connecting the predicted transmembrane helices are denoted *L1* to *L6*.

PM protein samples were incubated at 55 °C for 5 min prior to separation by 10% SDS-PAGE. The Va^{2+}-ATPase method was used to assess the degree of purity of the plasma membrane vesicles (26). Reactions were carried out in the presence or absence of 5 μM sodium orthovanadate (BDH). The percentage of Va^{2+}-dependent ATPase activity was calculated by 1 − [(Va^{2+}-dependent ATPase/total ATPase)].

To investigate the association of Mlo with plasma membranes, equivalent amounts of PM proteins were resuspended in lysis buffer, in 0.1 M Na_2CO_3 pH 11, in Triton X-100 or Sarkosyl at 1% final concentration. After incubation on ice for 30 min, the mixtures were centrifuged at 125,000 × *g* for 30 min. The presence of Mlo in the resulting pellet or supernatant fractions was determined by immunoblot assays with the MYP antibody as described above.

Computational Methods—DNA sequences were assembled and analyzed by using the University of Wisconsin Genetic Computer Group (GCG) computer packages (27) and the STADEN software package for UNIX users (fourth edition, 1994). *Arabidopsis thaliana* (AtDB), PIR, SWISSPROT, and GeneBank data bases were searched for sequence similarities by using the BLAST function. Related sequences were aligned with the programs PILEUP and CLUSTALW, and the degree of sequence relatedness was determined by using BESTFIT, GAP, DISTANCES, and GROWTREE and adjusted by eye when appropriate.

The deduced Mlo protein sequence was analyzed with the following prediction methods for polytopic membrane proteins, TMHMM (28), TOPPREDII (29), TMPRED, TMAP, and PHD.

The genome-wide analyses were performed as described previously (30). All predicted open reading frames in each genome were analyzed using TopPredII for the presence and orientation of membrane spanning segments. The results were presented as a contour plot with the number of membrane spanning segments and the length of the protein on the *y* and *x* axes, respectively. In order to reduce the number of soluble proteins falsely predicted as membrane proteins, the following selection criterion was adopted. At least two segments with an average hydrophobicity of at least 1.5 were selected for TopPredII calculations. The data base versions used were *Caenorhabditis elegans* Wormpep data base version 16 and *A. thaliana* data of May 1999. In all cases the open reading frames are predicted from the genomic sequence.

RESULTS

We initiated our membrane topology studies by analyzing the deduced Mlo sequence with a number of prediction methods for polytopic membrane proteins. Different membrane topologies were predicted by the various algorithms, suggesting the existence of 6 or 7 hydrophobic segments with the potential to form transmembrane helices. We present data from the transmembrane hidden Markov model algorithm (TMHMM) used to guide our membrane topology experiments, because it indicated the maximum number of potential hydrophobic transmembrane stretches in Mlo (Fig. 1, *upper panel*). TMHMM

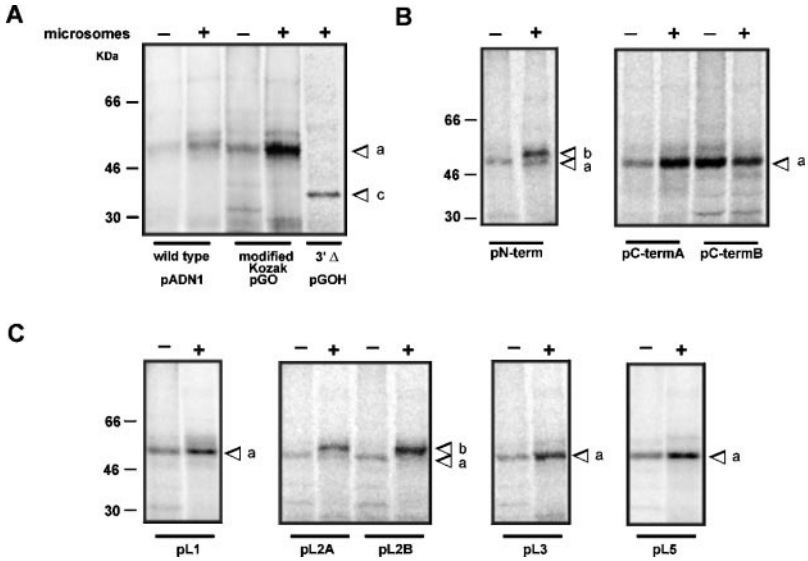

FIG. 2. **Scanning *N*-glycosylation mutagenesis of Mlo.** Autoradiographs of SDS-PAGE separated proteins following *in vitro* translation of mRNA obtained by *in vitro* transcription of plasmids containing Mlo variants. *In vitro* translation was carried out in the presence of L-[^{35}S]methionine. For each Mlo derivative, translations were carried out in the presence (+) or absence (−) of dog pancreas microsomes. Designations of each tested plasmid are indicated below the corresponding lanes of the autoradiographs. *A*, *in vitro* translation products obtained with plasmids containing wild type *Mlo* (*pADN1*), wild type *Mlo* containing the modified context 5′ to the initiator ATG codon (*pGO*), and a construct containing a 3′ truncated coding sequence (*pGOH*). *B*, *in vitro* translation products obtained with plasmids each containing single *N*-glycosylation acceptor sites in the N terminus (*pN-term*) and in the C terminus (*pC-termA* and *B*). *C*, *in vitro* translation products obtained with plasmids each containing single *N*-glycosylation acceptor sites in predicted loop 1 (*pL1*), loop 2 (*pL2A, pL2B*), loop 3 (*pL3*), and loop 5 (*pL5*). *Arrows* denote the position of non-glycosylated (*a*) and glycosylated (*b*) full-length wild type Mlo and full-length Mlo variants. The position of a C-terminal truncation Mlo derivative is denoted by *c*. *Numbers* on the *left* indicate the molecular mass (kDa) corresponding to ^{14}C-methylated proteins used as size markers.

predicts an extracellular N-terminal segment, 7 TM domains varying in length between 18 and 22 amino acids (denoted I–VII), and an intracellular C-terminal domain (Fig. 1, *lower panel*). The model also defines stretches for six interhelical loops, L1–L6.

Scanning N-Glycosylation Mutagenesis of Mlo—We used *N*-glycosylation mutagenesis to determine the number and the orientation of the predicted interhelical loops of Mlo relative to the membrane, because this method introduces minimal structural alterations in the wild type protein. Wild type Mlo lacks endogenous Asn-Xaa-Ser/Thr *N*-glycosylation acceptor sites, providing therefore an ideal template for this approach. We employed *in vitro* transcription/translation of wild type *Mlo* and a number of *Mlo* derivatives in a cell-free system consisting of a rabbit reticulocyte lysate supplemented with canine microsomal membranes. Site-directed mutagenesis was used to introduce single glycosylation acceptor sites (N-*X*-S/T) at various positions in the predicted Mlo loops. A lumenal or cytoplasmic location of a given glycosylation site was then inferred from the *N*-glycosylation status of wild type and Mlo derivatives. *N*-Glycosylation corresponding to the co-translational attachment of the high mannose oligosaccharide at a single acceptor site is expected to increase the molecular mass of Mlo by about 2.5 kDa. The glycosylation reaction can be catalyzed by the oligosaccharyl transferase (OST) only if the glycosylation acceptor site resides in a peptide loop exposed to the lumenal side of the microsomes (31).

Expression of wild type Mlo cDNA in the vector revealed a weak signal at approximately 52 kDa (Fig. 2*A*, *wild type*, *pADN1*). Modification of the sequence immediately 5′ to the initiator ATG codon of *Mlo* to a "Kozak consensus" resulted in a more than 10-fold increase in the translation rate of the 52-kDa species (plasmid pGO; Fig. 2*A*, modified Kozak), suitable for subsequent *N*-glycosylation scanning studies. Since the apparent molecular mass of the 52-kDa *in vitro* translation product deviated substantially from the predicted 60.4-kDa Mlo wild type protein, we tested the identity of the 52-kDa signal by *in vitro* translation of a construct containing a 3′ Mlo truncation derivative with an expected molecular mass of 43 kDa. Expression of the plasmid pGOH generated a signal with an apparent molecular mass of 35 kDa (Fig. 2*A*, *3′ del.*). The proportional 8-kDa difference between observed and predicted molecular masses obtained for both wild type Mlo and 3′ deletion constructs can be explained by their common predicted hydrophobic segments, leading to an accelerated mobility in SDS-PAGE. This is a characteristic feature of integral membrane proteins (32). All subsequent modifications of the coding region of *Mlo* for *N*-glycosylation analysis were based on the construct containing the Kozak consensus sequence (pGO).

We introduced single glycosylation acceptor sites in the N- and C-terminal segments and in putative loops 1, 2, 3, and 5 (Fig. 2, *B* and *C*). To minimize structural perturbations of the wild type topology, single or double amino acid substitutions were introduced in poorly conserved regions (deduced from a compilation of *Mlo* family member sequences; see below) and, where possible, residues with similar characteristics as the acceptor site ones were mutated. Introduction of a glycosylation acceptor site in the N terminus (pN-term) resulted in a higher molecular mass signal compared with wild type Mlo, while no change in mass was detected following *in vitro* translation of two plasmids containing C-terminal constructs (Fig. 2*B*, *pC-term A* and *pC-term B*). The shift in molecular mass of approximately 2.5 kDa was observed only in the presence of microsomes, indicating glycosylation and, therefore, a lumenal location of the N terminus. In contrast, two C-terminal derivatives, pC-term A and pC-term B, gave no change in mass and, indicating a cytoplasmic location for the C terminus.

A precise distance constraint allows glycosylation of acceptor sites only when the Asn residue is located at a minimum of 12–14 amino acid residues both upstream and downstream of a transmembrane segment (33, 34). Therefore, all acceptor sites in putative loops 1, 2, 3, and 5 were placed further away from any predicted flanking transmembrane segment. For this reason, we extended the size of loop 1 by inserting a 14-residue

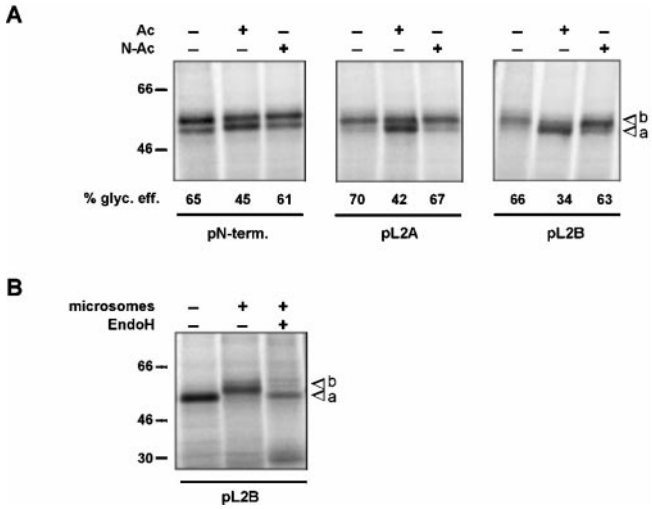

FIG. 3. **Competitive inhibition and enzymatic removal of *N*-glycosylation.** Autoradiographs of SDS-PAGE separated protein products following *in vitro* translation of mRNA obtained by *in vitro* transcription of plasmids containing Mlo variants. *A*, competitive inhibition of *N*-glycosylation by addition of acceptor (*Ac*) or non-acceptor (*N-Ac*) peptides. Competitive inhibition is shown for *in vitro* translation products obtained with plasmids pN-term, pL2A, and pL2B. Presence or absence of the inhibitors is denoted by (+) or (−). The efficiency of *N*-glycosylation (*glyc. eff.*) was calculated from the integrated areas corresponding to the *N*-glycosylated and non-glycosylated full-length products. The relative glycosylation efficiency (%) is indicated for each treatment below each lane. *B*, removal of *N*-linked oligosaccharides by Endo H treatment. *In vitro* translation product obtained with plasmid pL2B was treated with Endo H for 30 min after insertion into dog pancreatic microsomes. All other symbols are identical to those shown in Fig. 2.

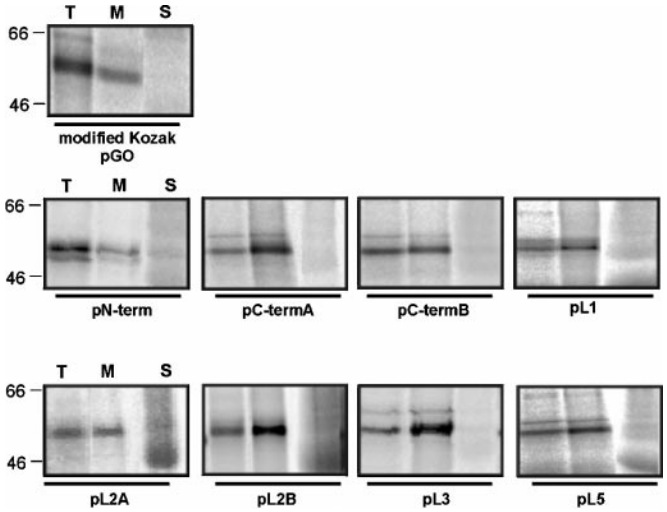

FIG. 4. **All Mlo variants integrate into dog pancreas microsomal membranes.** Figure shows autoradiographs of SDS-PAGE separated protein products following *in vitro* translation of mRNA obtained by *in vitro* transcription of plasmids containing Mlo variants, and alkaline extraction of *in vitro* translation products obtained with the indicated plasmids. Each crude translation mixture (*T*) is shown next to membrane precipitate (*M*) and supernatant (*S*) obtained after Na_2CO_3 (pH 11, 0.1 M) treatment. Variability in signal intensities, due to differences in the alkaline extraction efficiencies, can be observed for the glycosylated and non-glycosylated products in independent experiments.

peptide encompassing the acceptor site introduced into loop 2 (construct pL1). This alteration caused an expected increase in mass of 1.6 kDa relative to wild type *Mlo* but did not reveal a detectable glycosylation product upon addition of microsomes (Fig. 2*C*, *pL1*). In contrast, higher molecular band shifts were observed in a microsome-dependent manner for both pL2A and pL2B constructs, indicating efficient *N*-glycosylation of each introduced glycosylation acceptor site. Finally, no mobility shift was observed in the presence of microsomes following expression of derivatives expressed by constructs pL3 and pL5 (Fig. 2*C*). Taken together, the data indicate a lumenal orientation for the N terminus and loop 2, and a cytoplasmic orientation for loops 1, 3, and 5 and the C terminus.

All N-Glycosylated Mlo Variants Retain the Ability for Membrane Integration—To confirm that the observed shifts in Mlo mobility were the result of the catalytic activity of OST, we supplemented one translation reaction in each set of translation mixtures with a competitive inhibitor of *N*-glycosylation, the acceptor tripeptide *N*-benzoyl-Asn-Leu-Thr-*N*-methylamide (Ac). For the N terminus and the two loop 2 glycosylation variants (Asn^2, Asn^{104}, and Asn^{111}, respectively), glycosylation was significantly inhibited by the acceptor tripeptide (Fig. 3*A*). This is seen as a change in the ratio of glycosylated variants over non-glycosylated ones in the presence and absence of the competitive inhibitor (quantified as incorporated ^{35}S). In contrast, unaltered glycosylation efficiencies were observed upon addition of a non-acceptor control peptide *N*-benzoyl-Asn-Leu-(allo)Thr-*N*-methylamide (N-Ac). Additional evidence for a specific glycosylation was obtained by treatment of the loop 2 variant pL2B with endoglycosidase (Endo H), which removes *N*-linked oligosaccharides. A predicted mass reduction of 2.5 kDa was observed in presence of Endo H (Fig. 3*B*).

To ensure that all tested variants of Mlo retained the ability to integrate efficiently into dog pancreatic microsomal membranes, whether or not they were glycosylated, the *in vitro* translation for each construct was followed by extraction with 0.1 M Na_2CO_3. Both the *N*-glycosylated and non-glycosylated products were found to be resistant to alkaline extraction (Fig. 4).

The Major Portion of Mlo Is Exposed on the Cytoplasmic Side of the Membrane—Proteinase K and trypsin digestions were used as an indirect approach to localize Mlo with respect to the membrane. Cytoplasmic domains remain outside of the microsomes, rendering them sensitive to these enzymes. Only those domains that are translocated into the microsomal lumen and/or reside in the membrane are recovered in the pellet after alkaline extraction. Fig. 5 shows the results obtained from digestion of the *in vitro* translated protein products of Mlo (pGO; *lanes 1–3*) and a loop 2 variant (pL2B; *lanes 4–6*). A major signal corresponding to a fragment of an apparent mass of approximately 16 kDa was detected after proteinase K digestion and of approximately 18 kDa after trypsin treatment (*lanes 2* and *3*, respectively). The larger size of the fragments released by protease digestions from variant pL2B (*lanes 4–6*) is consistent with an *N*-glycosylation of loop 2. The protease resistant fragment thus includes loop 2, and probably represents an N-terminal portion of Mlo resulting from cleavage within loop 3. The absence of protease-resistant fragments larger than 16/18 kDa suggests that the major portion of Mlo is located on the cytoplasmic side of the membrane.

Hydrophobic Segments I–VI Have Membrane Spanning Properties—Loops 4 and 6 had predicted sizes of 3 and 22 amino acids, respectively. Since long amino acid extensions in these loops would have unknown effects on the protein topology and were required to introduce *N*-glycosylation acceptor sites accessible to the OST, we chose a different approach. This approach enabled us to test the membrane spanning properties of Mlo HS I–VI.

We made fusions between various parts of Mlo and Lep, an *E. coli* inner membrane protein (35). Lep was previously used in numerous studies of membrane protein topology, both in *E. coli* and in the microsomal *in vitro* system, and it was used as a scaffold for presenting potential Asn-Ser-Thr acceptor sites to

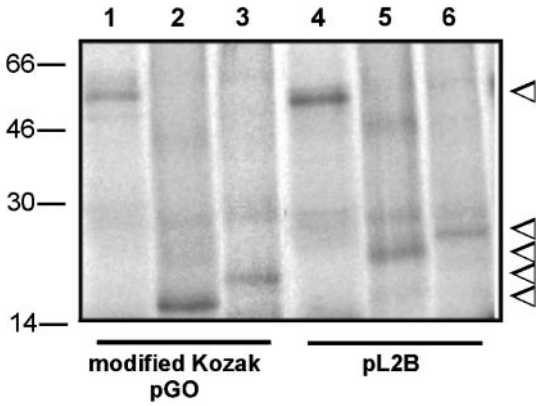

FIG. 5. **Protease digestion of Mlo integrated in dog pancreas membrane.** Protease digestions of *in vitro* translation products obtained with plasmids pGO and pL2B were carried out with proteinase K and trypsin. All reactions were performed in presence of microsomes and were followed by alkaline extractions. All products were separated by a 12% SDS-PAGE. *Lanes 1–3*, *in vitro* translation of RNA obtained with plasmid pGO. *Lanes 4–6*, *in vitro* translation of RNA obtained with plasmid pL2B. *Lanes 1* and *4*, alkaline extraction only; *lanes 2* and *5*, proteinase K digestion followed by alkaline extraction; *lanes 3* and *6*, trypsin digestion followed by alkaline extraction. *Arrows* indicate the position of the major signals obtained for each treatment in independent experiments.

the OST (17, 33, 34, 36). Lep has two transmembrane segments (H1 and H2), which are connected by the cytoplasmic loop P1. Engineered sites in the coding region of *Lep* allowed the insertion of different *Mlo* fragments generating Lep-Mlo-Lep sandwich constructs. In addition, we replaced *Lep* domains with Mlo fragments, generating Lep-Mlo fusions. Translocation of the P2 domain into the microsomal lumen is easily detected by the modification of a potential *N*-glycosylation acceptor site that is present in the wild type sequence (17).

HS I–VI were tested for their ability to insert into the endoplasmic reticulum membrane either as isolated segments or in the context of Mlo C-terminal truncation mutants. Fusion constructs were obtained by substituting the second transmembrane domain (H2) of Lep (plasmids pB5N6 and pB1N4) and by fusing the C-terminal domain (P2) of Lep downstream of a 3′ truncated Mlo (pXN4). In addition, we fused the H2 and the P2 domains of Lep downstream of a 3′ truncated Mlo (pXB5; Fig. 6). The membrane topology of the four tested Mlo-Lep fusion proteins was inferred from the glycosylation of the P2 domain. All four *in vitro* translated constructs were efficiently glycosylated in the presence of microsomes (Fig. 6). Glycosylation of the protein product derived from construct pB5N6 demonstrates that HS VI translocates across the membrane. Glycosylation of the fusion proteins corresponding to constructs pB1N4, pXN4, and pXB5 is consistent with the formation of 3, 4, and 5 transmembrane segments, respectively, in different Lep contexts. In each case, the specificity of *N*-glycosylation was confirmed by competition experiments with the acceptor tripeptide (data not shown). We conclude that HS I–VI of Mlo all form transmembrane segments. Together with the results described above for the N- and C-terminal domains and loops 1, 2, 3, and 5, these observations provide strong support for the predicted topology shown in Fig. 1.

Subcellular Localization of Mlo—Subcellular localization of Mlo in barley leaf tissue was studied by differential centrifugation of cell extracts, separating membrane vesicles by mass and size. Fractions were tested for the presence of Mlo by immunoblotting with an antibody (designated MYP), raised against a C-terminal Mlo peptide (see below). Mlo was undetectable in soluble or 10,000 x *g*, 20,000 x *g*, and 50,000 x *g* microsomal fractions (Fig. 7). A weak signal was detected in the 100,000 x *g* high speed pellet corresponding to a mass of approximately 50 kDa. In contrast, antibodies against marker proteins for the ER (calreticulin) and the plasma membrane H^+-ATPase revealed signals of the expected mass mostly in the lower speed (10,000 x *g* and 20,000 x *g*) and high speed pellet (50,000 x *g* and 100,000 x *g*) fractions, respectively (Fig. 7). This indicated a clear separation between ER and plasma membrane vesicles.

The weak signal detected by the MYP Mlo antibody in the 100,000 x *g* high speed H^+-ATPase-rich pellet prompted us to further enrich PM vesicles by using the two-phase partitioning methodology (25). In this highly enriched PM fraction, judged by the percentage of vanadate-dependent ATPase activity above 80%, a strong signal was detected with the MYP antibody with an apparent mass of 52 kDa (Fig. 7, *lane PM*). Densitometric analysis revealed that the H^+ATPase marker protein was enriched 10 times, the calreticulin 1.5 times, and the Mlo protein 25 times in the PM fraction, when compared with the strongest signal in other cellular membrane fractions.

To confirm that the MYP antibody specifically recognized Mlo, immunoblot analysis and immunoprecipitation were carried out using aliquots of *in vitro* translated Mlo (Fig. 8*A* and data not shown). *In vitro* translated Mlo was recognized in a dose-dependent manner by the MYP antibody, and the mass of the detected signal was found to be comparable to that observed *in vivo* in highly enriched PM fractions (Fig. 8*A*). The slight difference in molecular mass between *in vitro* translated and PM isolated Mlo protein could be due to the delipidation treatment of the PM vesicles or other post-translational modifications known to occur in plant cells (*e.g.* phosphorylation, palmitoylation).

Further evidence for the specificity of the MYP antibody was obtained by comparing PM fractions of *Mlo* wild type and two *mlo* deletion mutants (*mlo-3* and *mlo-4*). In both mutants, the deletions (2 base pairs in *mlo-3* and 11 base pairs in *mlo-4*) cause frameshifts, truncating the C-terminal part of the wild type protein (9). Immunoblot analysis with the MYP antibody revealed complete absence of the 52-kDa wild type signal in both *mlo-3* and *mlo-4* mutants (Fig. 8*B*). This supports the specificity of the MYP antibody as the MYP peptide encompasses the region truncated in both deletion mutants.

Mlo Is an Integral Membrane Protein—We tested different protein extraction methods of leaf PM fractions for their ability to solubilize Mlo. Neither alkaline treatment, known to be effective in extracting proteins that are peripherally associated with membranes, nor addition of the non-ionic detergent Triton X-100 (37) was able to extract Mlo from the plasma membrane vesicles (Fig. 9, *lanes 1–4*). However, we removed Mlo from the pelleted leaf PM vesicles and recovered the solubilized protein in the supernatant fraction upon treatment with the ionic detergent Sarkosyl, which affects the integrity of the lipid bilayer (Fig. 9, *lanes 5* and *6*). This provides direct evidence that Mlo is an integral polytopic membrane protein.

The 7 TM Structure of Mlo Is Conserved in All Tested Family Members—*Mlo* sequence-related genes were found both in monocot and dicot plant species. In *A. thaliana* we identified 17 family members, of which 12 are full-length and are distributed on different chromosomes (Table I; data derived from genomic and EST data bases). Similarly, nine homologues were found in the *Zea mais* EST data base.[3] ESTs with significant sequence-relatedness to Mlo were also identified in *Linum usitatissimum*, *Populus balsamifera*, *Glycine max*, *Brassica rapa*, *Gossypium hirsutum*, and *Lycopersicon esculentum* (Table I). Since the current *Arabidopsis* genomic data base repre-

[3] J. Duvick, personal communication.

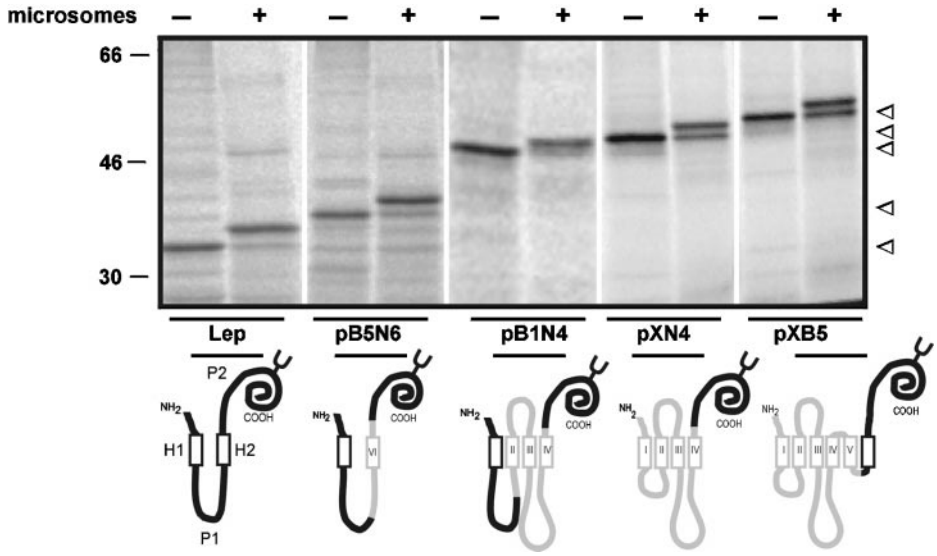

FIG. 6. ***N*-Glycosylation analysis of Lep-Mlo fusion proteins.** Autoradiographs of SDS-PAGE separated protein products following *in vitro* translation of mRNA obtained by *in vitro* transcription of plasmids containing Lep-Mlo fusions. Presence or absence of microsomes is denoted by (+) and (−), respectively. Schematic representations of the deduced topology for wild type Lep and each Lep-Mlo fusion construct are shown below each tested plasmid. Lep domains are indicated by a *black line*, and Mlo domains are denoted by a *gray line*. *Roman numerals* refer to the hydrophobic segments shown in Fig. 1. *P1* and *P2* denote the cytoplasmic and C terminus of Lep. *H1* and *H2* denote the two transmembrane helices of Lep. *Y* indicates the position of the single *N*-glycosylation site in the C terminus of Lep. *Black arrows* indicate the position of non-glycosylated translation product obtained with each tested plasmid.

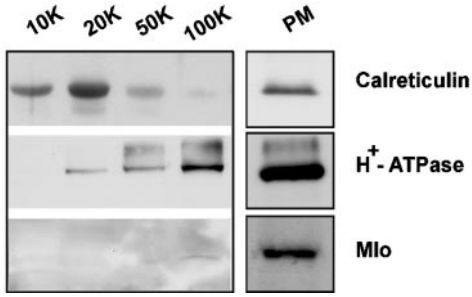

FIG. 7. **Subcellular localization of Mlo.** Immunoblot analysis of barley leaf microsomal fractions probed with antibodies specific for Mlo, the ER marker protein calreticulin or the plasma membrane marker protein H^+-ATPase. Total leaf microsomes were separated by differential centrifugation into 10,000, 20,000, 50,000, and 100,000 x *g* fractions. PM vesicles were enriched from total leaf microsomes by using the two-phase partitioning method (25). Protein extracts of each fraction were separated by 10% SDS-PAGE. The percentage of the Va^{2+}-dependent ATPase activity in the PM fraction is 85%.

sents approximately 52% of the *Arabidopsis* genome (April 1999), we estimate the total number of *Mlo* homologues will total 25–35. Individual analysis of these sequence-related genes by the TMHMM algorithm revealed for both monocot and dicot family members almost identical hydrophobic segment predictions (data not shown). Since these predictions match the 7 TM topology determined experimentally for barley Mlo, we conclude that all available Mlo family members are likely to share the same membrane topology.

The common topology of the sequence-related genes enabled us to identify domains exhibiting either a high degree of sequence conservation or sequence variability (Fig. 10; calculated on the basis of 20 deduced full-length and 15 partial sequences). On average, family members exhibit 45% identity and 70% similarity. The most closely related Mlo family members identified so far are barley Mlo2, rice OsMlo1, and *Arabidopsis* AtMlo6 and AtMlo18. They share at amino acid level similarities ranging between 67 and 86%. The most distantly related family member, AtMlo4, exhibits 31% identity and 55% similarity with barley Mlo.

A striking sequence conservation was identified in each of the 7 TM segments, in cytoplasmic loops 1, 2, and 3, and the first 30 residues of the C terminus adjacent to TM 7 (Fig. 10). Highly variable sequence stretches are confined to extracellular loops 1 and 3, the N terminus, and C-terminal sequences distal to TM 7. It is notable that four cysteine residues, located in the highly variable extracellular loops 1 and 4, are strictly conserved among all family members.

Mlo Represents the Single Most Abundant Protein Family with a 7 TM Topology in Arabidopsis—The large number of family members identified in *Arabidopsis* and maize prompted us to perform a statistical genome-wide analysis of multi-spanning integral membrane proteins in *Arabidopsis* (Fig. 11). The contour plots reveal the frequency of proteins with a given number of predicted transmembrane segments and amino acid length for the genomes of *A. thaliana*, and for comparative purposes, the metazoan *C. elegans* genome. In *Arabidopsis*, abundant classes are represented by proteins containing 6 and 12 TMs. These polytopic membrane proteins have a length of 275 ± 50 and 500 ± 50 amino acids, respectively. A close inspection of the *Arabidopsis* data base identified *Mlo* as the only gene family encoding 7 TM proteins, corresponding to approximately 0.3% of the genome. In contrast, the major class in *C. elegans* is represented by 7 TM proteins of approximately 325 ± 50 amino acids, corresponding to approximately 5% of the total genome.

DISCUSSION

A Common 7 TM Scaffold for the Mlo Family—Different orientations and numbers of TM domains are predicted for Mlo when different computational methods were employed for structural analysis of the deduced protein sequence. Here we have shown experimental evidence for a 7 TM topology of Mlo, using a combination of scanning *N*-glycosylation mutagenesis and Lep-Mlo fusion proteins.

A major difficulty of computational methods is to resolve whether two closely linked hydrophobic protein stretches form a single or two separate TM helices. This problem is illustrated by helices IV and V of Mlo, which are connected by the very short extracellular loop 2. Because loops shorter than 26 amino acids are generally inaccessible for the OST catalytic site (33, 34), we had to test the membrane spanning potential of hydrophobic segment IV by using a Lep-Mlo-Lep sandwich and a Lep-Mlo fusion construct. Both constructs demonstrated the membrane spanning properties of the 17 amino acid hydrophobic segment IV by driving the C terminus of Lep into the ER lumen. Together with the results derived from two constructs designed to test the membrane spanning potential of the neighboring hydrophobic segment V, we conclude that hydrophobic stretches IV and V form two rather than one TM helix.

The extremely short connecting loop between helices IV and V (three amino acids according to TMHMM) may impose a "hairpin" structure in the segment comprising helix IV-loop 4-helix V, possibly enforcing inter-TM helical interactions. The

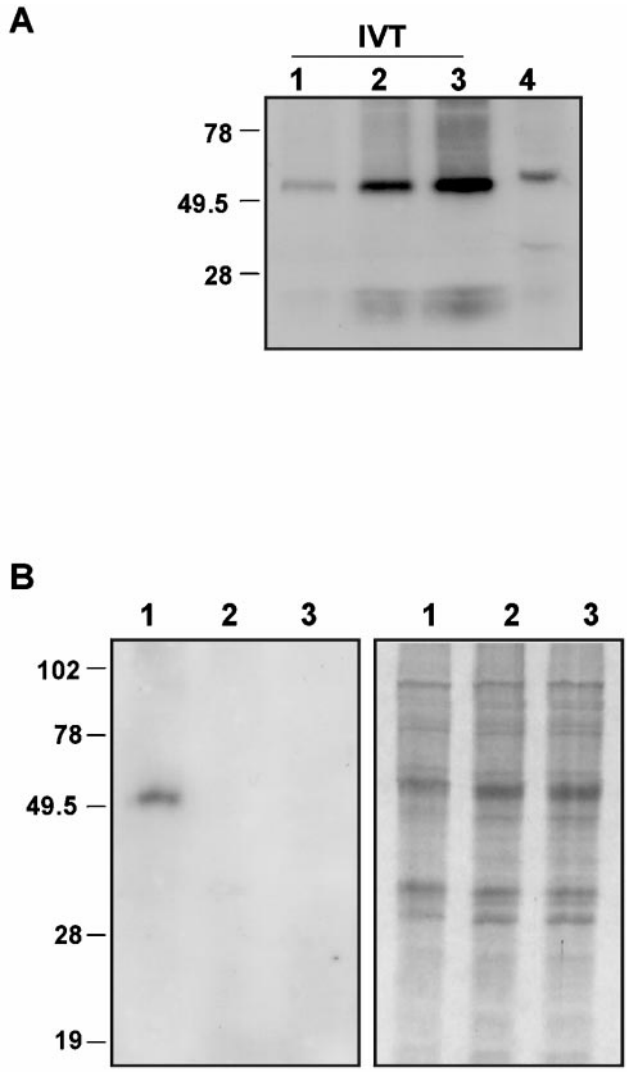

FIG. 8. **The MYP antibody recognizes barley Mlo.** *A*, Immunoblot analysis of *in vitro* translated Mlo protein and leaf PM vesicles purified from *Mlo* containing cultivar Ingrid. 1, 2, and 5 μl of *in vitro* translation reactions (*IVT*) obtained with plasmid pGO were separated by 10% SDS-PAGE (*lanes 1*, *2*, and *3*). 50 μg of PM proteins were separated in *lane 4*. *Numbers* on the *left* indicate the molecular mass (kDa) corresponding to pre-stained broad range size marker proteins (Life Technologies, Inc.). *B*, immunoblot analysis (*left*) and corresponding Coomassie-stained SDS-PAGE (*right*) of plasma membrane fractions isolated from *Mlo* wild type (*lane 1*), *mlo-3* (*lane 2*), and *mlo-4* (*lane 3*) genotypes. Enrichment of plasma membrane vesicles was calculated as percentage of Va^{2+}-dependent ATPase activity. *Mlo*, 85%; *mlo-3*, 88%; *mlo-4*, 79%.

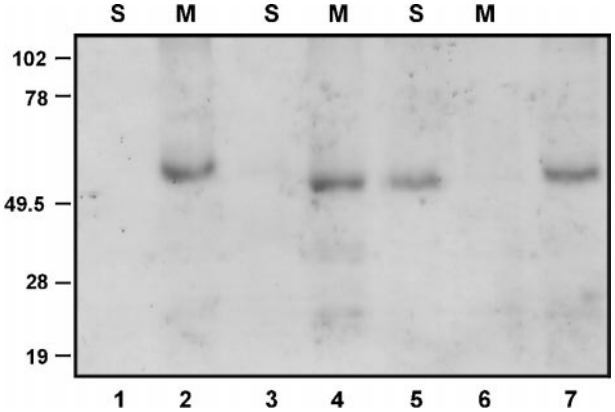

FIG. 9. **Mlo is an integral membrane protein.** Western blot analysis of protein extracts derived from purified leaf plasma membrane vesicles of the *Mlo* containing genotype Ingrid. Soluble (*S*) and membrane (*M*) protein fractions were obtained subsequent to 125,000 × *g* centrifugation following treatment with 0.1 M Na_2CO_3 pH 11 (*lanes 1* and *2*), 1% Triton X-100 (*lanes 3* and *4*), or 1% Sarkosyl (*lanes 5* and *6*). Total plasma membrane proteins extracted with lysis buffer only were separated in *lane 7*. Equal amounts of proteins (50 μg) were loaded in each lane. *Numbers* on the *left* indicate the molecular mass (kDa) corresponding to pre-stained broad range size marker proteins (Life Technologies, Inc.).

clear distance conservation between helices IV/V among family members (3–5 amino acids) identifies the hairpin structure as a topological signature of the Mlo family. Interestingly, our attempts to probe loop 4 by insertion of a 39-amino acid peptide containing a GAS failed due to poor membrane insertion properties of the resulting derivative. Possibly, the hairpin structure of helix IV/V may be critical for efficient membrane insertion of Mlo.

Analysis of the currently available 20 full-length Mlo family member sequences by TMHMM predicts in each case 7 hydrophobic segments at similar positions relative to barley Mlo (data not shown). Lep-Mlo fusion construct pB5N6 allowed us to clearly identify the existence of transmembrane helix VI for the barley Mlo whose presence was controversial on the basis of previous computational analysis using a single family member sequence (9). Thus, computational analysis of sequences involving multiple family members together with the experimental topology studies suggest the existence of a common 7 TM scaffold. TMHMM calculations on the basis of this scaffold predict that on average 15% of a Mlo type protein is exposed extracellular, 25% is buried within the plasma membrane, and 60% is located within the cytoplasm (Fig. 10). These predictions are corroborated by protease digestion patterns of barley Mlo inserted into canine microsomes where a single large fragment of approximately 16–18 kDa resists degradation, indicating that the largest part of Mlo is likely to be cytoplasmic. The 7 TM scaffold also enabled us to identify conserved and variable domains among family members. Peptide sequences comprising each of the 7 TM helices and cytoplasmic loops 1, 2, and 3 are conserved among family members. In contrast, all extracellular domains and the cytoplasmic C-terminal tail are highly variable. Interestingly, the N- and C-terminal sequences and extracellular loop 1 are the only regions exhibiting variability in length among family members (differences range between 33, 93, and 47 amino acids, respectively).

The high degree of sequence conservation detected in each of the 7 TM domains of the Mlo family may simply reflect the limited choice of hydrophobic amino acids with a potential to constitute building blocks of a TM helix (38). However, conserved TM stretches are interspersed with highly variable residues in each helix. A similar distribution of conserved and variable residues is observed in cytoplasmic loops 1, 2, and 3. Thus, each of the conserved domains of the Mlo family are characterized by a patchwork of conserved stretches interspersed with variable amino acid residues. If one assumes that the *Mlo* family originated from a single ancestral gene, then the clear distinction between highly variable and conserved domains among family members provides strong evidence for domain-specific diversifying selection.

A notable feature of the sequence-variable extracellular loops 1 and 3 are four invariant cysteine residues. This strongly suggests their involvement in the formation of disulfide bridges within the same and/or between extracellular loops. Extracellular inter-loop disulfide bridges are frequently found in mammalian 7 TM receptors and are thought to be critical in stabilizing the relative arrangement of TM helices to each other (39, 40). The conservation of the extracellular cysteine residues among Mlo family members is evidence for a common higher order structure.

TABLE II
Primer sequences used to obtain Mlo derivatives by SOE

Plasmid	Primer		Mutations introduced in Mlo
pN-term	NTERMs	5′-GTG TCG GGC TCG ACC TCC GAC AAA AAA GGG GTG C-3′	Asn and Val inserted after Met[1] and Gly-Ser-Thr-Ser after Ser[2]
	3595a	5′-GGT GAT GAC GCT GTA G-3′	
	NTERMa	5′-GTC GAG CCC GAC ACG TTC ATG GTG GCA AGC TTG -3′	
	2750s	5′-ACA TGT TCT TTC CTG -3′	
pL2A	EC1as	5′-ACC AAC GGC TCC AAG CCC AGC AAG TAC -3′	$Glu^{104} \rightarrow Asp$
	4353a	5′-GAC CTC TTC ATG TTT G -3′	$Arg^{106} \rightarrow Ser$
	EC1an	5′-CTT GGA GCC GTT GGT GCC GCG CTT G -3′	
	2970s	5′-CGC AAC GCA ATT AAT G -3′	
pL2B	EC1bs	5′-AAG AAC GTT ACC TAC TGC CCG GAG G -3′	$Tyr^{111} \rightarrow Asn$
	4353a	5′-GAC CTC TTC ATG TTT G -3′	$Asp^{113} \rightarrow Thr.$
	EC1ba	5′-GTA GGT AAC GTT CTT GCT GGG CTT GC -3′	
	2970s	5′-CGC AAC GCA ATT AAT G -3′	
pL3	IC2s	5′-TTC AGG AAC GTC ACC AAG GTG GAC -3′	$Ser^{214} \rightarrow Asn$
	4353a	5′-GAC CTC TTC ATG TTT G -3′	
	IC2a	5′-CTT GGT GAC GTT CCT GAA GAA CTG C -3′	
	2970s	5′-CGC AAC GCA ATT AAT G -3′	
pL5	IC3s	5′-CGG ACC AGC TCC ATC AAG GGG GCC CCC GTG -3′	$Ala^{313} \rightarrow Asn$
	4405a	5′-CCT TGG ACG TCT GCT CGT C -3′	$Val^{315} \rightarrow Ser$
	IC3a	5′-GAT GGA GCT GGT CCG GTC CTG GAT CTC -3′	
	3321s	5′-CTG GAG AAG ATG AAG GCG -3′	
pC termA	CTERMas	5′-ATG AAC CGG TCG GAC GAC CCC CAG -3′	$Gly^{472} \rightarrow Asn$
	316a	5′-CAA CAC TCA ACC CTA TCT -3′	
	CTERMaa	5′-TCC GAC CGG TTC ATG CCC TTG TGA AGC -3′	
	4204s	5′-ATG GCG CAT TTT GTG TGG -3′	
pC-termB	CTERMbs	5′-ACA GGA GGA GGA ACG CCT C -3′	$Ser^{513} \rightarrow Asn$
	316a	5′-CAA CAC TCA ACC CTA TCT -3′	
	CTERMba	5′-AGG CGT TCC TCC TCC TGT C -3′	
	4204s	5′-ATG GCG CAT TTT GTG TGG -3′	
pL1	3311s	5′-GCC CGG AGG GCA AGG TGG CGC TCA TGT CCT GGG AGG CGC TGG AG -3′	14-amino acid insertion in loop 1 corresponding to the peptide in pL2B starting from Asn^{111} to Ser^{127}
	4351a	5′-ATG GAC CTC TTC ATG TTT G -3′	
	3310a	5′-CAC CTT GCC CTC CGG GCA GTA GGT AAC GTT CAG GGC CTT CTT GTG -3′	
	2972s	5′-CGC AAC GCA ATT AAT GTG -3′	

The gene structure of *Mlo* homologues is highly conserved (Table I). This reveals similar numbers of introns (11–14) as well as conserved intron/exon junctions. A striking observation is a single large exon encoding the cytoplasmic C-terminal tail of Mlo family members. The 5′ end of this exon corresponds precisely to the predicted junction linking the cytoplasmic C-terminal tail with TMVII in each family member. Since the cytoplasmic C-terminal tails are highly variable both in sequence and length, C-terminal variability of Mlo family members may have been generated by an exon shuffling process. It will be interesting to find out whether the variable cytoplasmic C termini are determinants contributing to functional specificity of individual family members.

Biogenesis and Subcellular Location of Mlo—A critical step in the biosynthesis of membrane proteins is their insertion into the lipid bilayer. This can either occur by sequential integration of individual TMs or by simultaneous insertion of multiple TMs. Our data provide evidence for the former. The hydrophobic segment VI alone is sufficient to force the C terminus of Lep into the ER lumen. This shows that the TM VI fragment functions as an independent unit containing information for directional membrane insertion. The results obtained with the other three Mlo-Lep fusion proteins demonstrated that various combinations of Mlo TM helices can each enforce the predicted lumenal translocation of the Lep C terminus. This suggests that they function independently from each other as topogenic signals. These data are consistent with a model in which membrane insertion is directed by a series of alternating signal anchor and stop transfer sequences (41, 42). However, we cannot yet discriminate whether Mlo membrane insertion occurs co- or post-translationally.

Both differential centrifugation of microsomes and selective enrichment of PM vesicles by phase partitioning provide clear evidence for a plasma membrane location of Mlo within leaf cells. The observed partial enrichment of the ER calreticulin marker protein in PM protein extracts was expected, given the continuity of ER and PM membranes within a cell and the documented enclosure of small ER vesicles within the larger plasma membrane vesicles (25). No MYP antibody signal was detected in the 20,000 x *g* vesicle-enriched fraction, consistent with the observation that Mlo does not contain an ER retention signal. Lack of ER retention signals in all currently available Mlo family members suggests that plasma membrane targeting is likely to be a general feature of the protein family. The PM localization of Mlo rules out previous speculation about a nuclear localization based on a putative nuclear localization sequence in the C terminus (9). Moreover, the sequence defining the putative nuclear localization sequence is not conserved among the family members described here.

Our pancreas microsome assays proved that efficient insertion of Mlo occurs in ER membranes, consistent with current data in eukaryotes indicating biosynthesis of plasma membrane proteins in the ER and subsequent transport through Golgi stacks into the plasma membrane (43–45). Signal peptides control the entry of virtually all proteins to the secretory pathway, both in eukaryotes and prokaryotes (46). They are usually located within the N-terminal part of the amino acid chain and can be cleaved off while the protein is translocated through the membrane (47). With two exceptions, AtMlo5 and AtMlo11, we identified cleavable or uncleavable signal peptides in all available Mlo family members at similar positions close to the N terminus, using the Signal P algorithm (47). This sug-

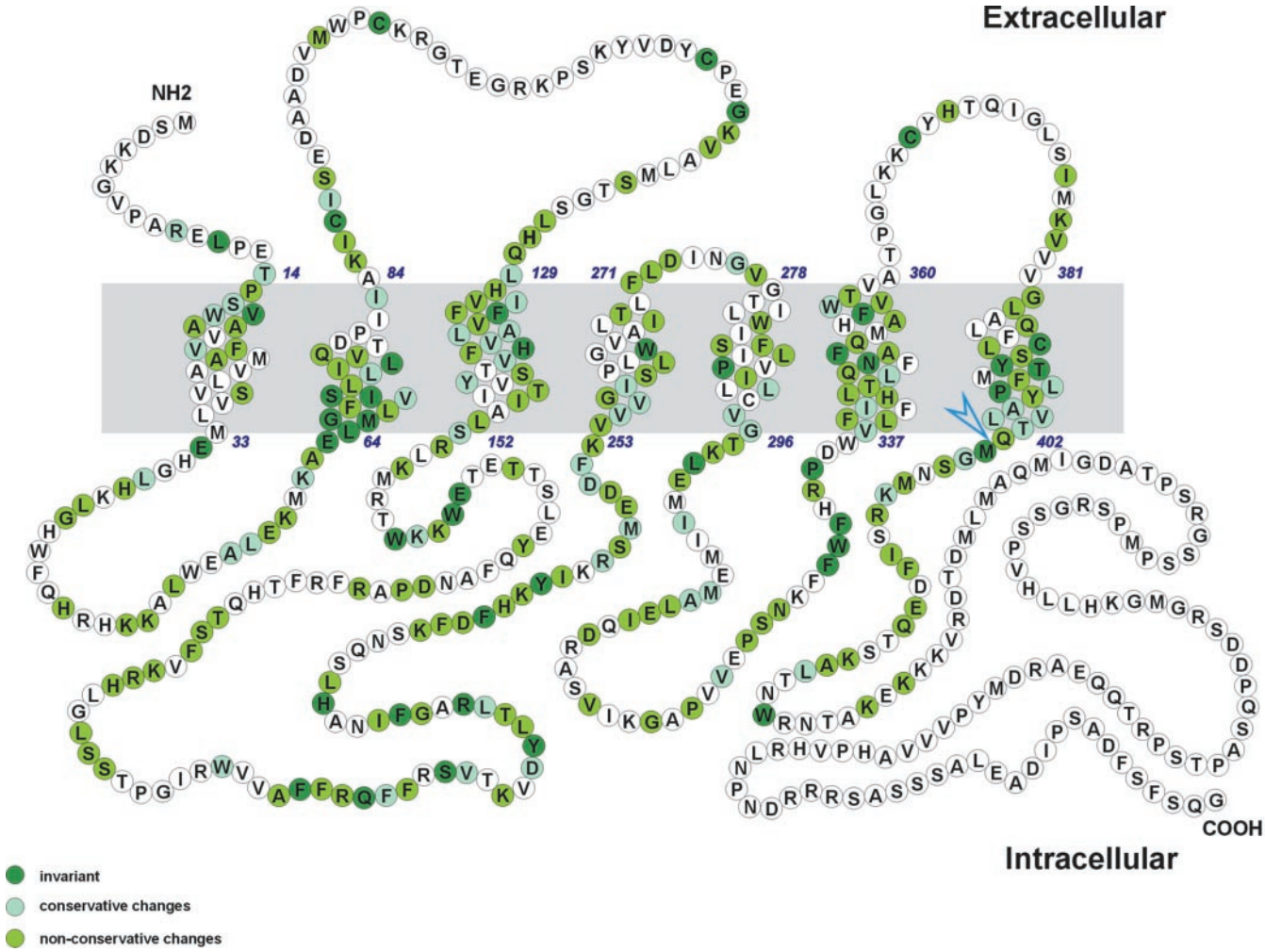

FIG. 10. **A common scaffold topology for the Mlo family.** Graphic representation of the deduced 7 TM topology of barley Mlo. The lipid plasma membrane bilayer is represented by a *gray horizontal bar*. *Circles* with *letters* represent amino acids identified by the *single-letter* amino acid code. Borders of the 7 TM helices were calculated using the TMHMM and TOPPRED II algorithms. Alignment of protein sequences from 20 full-length Mlo family members revealed highly variable and conserved residues (color-coded). Invariant residues, *dark green*; conservative changes in positions with at least 50% identical residues, *light green*; non-conservative changes in positions with at least 50% identical residues, *green*. *Numbers* indicate amino acid positions. The *arrow* indicates the position corresponding to the 5′end of exon 12.

gests that most family members are transported to the plasma membrane via a signal peptide-dependent secretory pathway. Our failure to detect Mlo in ER-enriched membrane fractions could be explained by a rapid transport to and a low turnover within the PM.

Mlo Uncovers the Only Abundant 7 TM Family in Higher Plants—The genetic diversification of *Mlo* family members within a single plant species, their topology, and subcellular localization are reminiscent of the most abundant class of 7 TM receptors in metazoans, the G-protein-coupled receptors (GPCRs) (48–50). GPCRs constitute a superfamily of 7 TM receptors that receive a signal at the extracellular side in the form of light or a chemical effector, which is then transduced through the membrane to the cytosolic side. The receptors relay extracellular signals through ligand binding into amplified intracellular responses via activation of the α subunit of heterotrimeric G proteins. The G protein in turn activates a second messenger system that elicits a response (51). Among membrane-bound receptors, GPCRs are the most diverse. There are at least five families of GPCRs showing negligible or no sequence similarity (52, 53).

Recently, a putative plant GPCR, GCR1, has been isolated in *Arabidopsis*. Although the topology of GCR1 has not been tested experimentally, its deduced membrane orientation and few sequence signatures are shared with mammalian GPCRs (54, 55). Silenced GCR1 plants exhibit a *Dainty* phenotype and reduced sensitivity to the plant hormone cytokinin benzyladenin, indicative for a role of GCR1 in cytokinin signaling. Contrary to the *Mlo* gene family, GCR1 is a single copy gene and none of the *Arabidopsis* Mlo family members shares significant sequence similarity with GCR1 or any of the GPCRs identified so far. A data base search in *A. thaliana* with members of each known GPCR family identifies only the single copy gene *GCR1*. This shows that plants lack known GPCR families and raises the question whether they have invented novel class(es) of GPCRs.

Our genome-wide analysis of multi-spanning integral membrane proteins in *A. thaliana* identifies the Mlo family as the only abundant class of 7 TM proteins. This analysis is limited to a current coverage of at least half of the *Arabidopsis* genomic sequence and a total of 25,000 *Arabidopsis* ESTs (56). The comparison of multi-spanning integral membrane proteins in *A. thaliana* and *C. elegans* underlines the prevalence of 7 TM proteins in the latter organism, representing approximately 5% of the gene complement. We estimate the total number of *Arabidopsis Mlo* homologues in the order of 25–35 family members. Since the total number of *Arabidopsis* genes is in the order of 20–25,000 (57, 58), the relative abundance of 7 TM

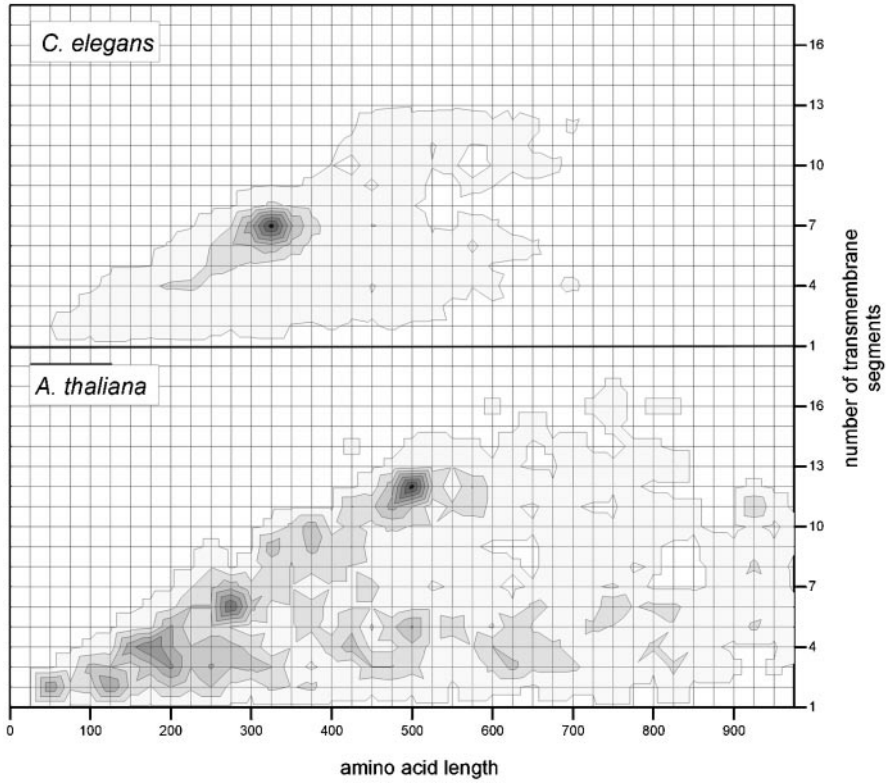

FIG. 11. **Genome-wide analysis of polytopic membrane proteins in *C. elegans* and *A. thaliana*.** Contour plots showing the frequency of proteins with a given number of predicted transmembrane helices and a given overall amino acid length in the genomes of *C. elegans* and *A. thaliana*. Computational analysis was carried out in bins of 25 residues (30). *Black dots* denote the highest frequency value in each panel and progressively *lighter shades* of *gray* indicate lower frequencies. For both organisms, each contour level represents a change in frequency by 12.5% of the maximum. The maximum number of proteins in any data point is 28 for *A. thaliana* and 264 for *C. elegans*.

proteins is at least an order of magnitude lower in comparison to the nematode *C. elegans*.

It is likely that *Arabidopsis* contains only a single Gα subunit, designated GPα1 (59). If *Mlo* homologues represent a novel class of GPCRs, then GPα1 serves as a candidate cytoplasmic interactor. Since GCR1 is also a candidate plant GPCR unrelated to the Mlo family, one would have to assume that GPα1 has the capacity to interact with unrelated receptors. Indeed, members of unrelated mammalian GPCR subfamilies can exhibit the same Gα coupling profile (60). One explanation is that receptor conformation is more critical for Gα selectivity and activation than receptor sequence conservation (53). Clearly, biochemical evidence is now needed to test the hypothesis whether *Mlo* homologues and GCR1 require for their function GPα1. The recent development of a single-cell transient expression system to study Mlo function provides a powerful tool to test the biochemical hypothesis proposed here *in vivo* (10).

Acknowledgments—We thank our colleagues Desmond Bradley, Tina Romeis, Ken Shirasu, and James Orme at the Sainsbury Laboratory for helpful suggestions on the manuscript. Alan Cavill, Margaret Corbitt, Patrick Bovill, and David Baker are acknowledged for their precious technical assistance. We are grateful to Dr. R. Napier and Dr. R. Serrano for providing the anti-maize calreticulin antibody and the 759 PM ATPase antibody, respectively.

REFERENCES

1. Baker, B., Zambryski, P., Staskawicz, B., and Dinesh-Kumar, S. P. (1997) *Science* **276,** 726–733
2. Jones, D. A., and Jones, D. G. (1997) *Adv. Bot. Res.* **24,** 90–167
3. Dietrich, R. A., Delaney, T. P., Ukness, S. J., Ward, E. R., Ryals, J. A., and Dangl, J. L. (1994) *Cell* **77,** 565–577
4. Frye, C. A., and Innes, R. W. (1998) *Plant Cell* **10,** 947–956
5. Arase, S., Fukuyama, R., Tokizawa, K., Ikegami, S., Honda, Y., and Nozu, M. (1997) *J. Phytopathol. Phytopatholog. Z.* **145,** 31–35
6. Freialdenhoven, A., Peterhänsel, C., Kurth, J., Kreuzaler, F., and Schulze-Lefert, P. (1996) *Plant Cell* **8,** 5–14
7. Aarts, N., Metz, M., Holub, E., Staskawicz, B., Daniels, M., and Parker, J. E. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 10306–10311
8. Jabs, T., Dietrich, R. A., and Dangl, J. L. (1996) *Plant Physiol.* **111,** 76
9. Büschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., van Daelen, R., van der Lee, T., Diergaarde, P., Groenendijk, J., Töpsch, S., Vos, P., Salamini, F., and Schulze-Lefert, P. (1997) *Cell* **88,** 695–705
10. Shirasu, K., Nielsen, K., Piffanelli, P., Oliver, R., and Schulze-Lefert, P. (1999) *Plant J.* **17,** 293–299
11. Erickson, B. W., and Merrifield, R. B. (1976) in *The Proteins* (Hill, R. L., and Neurath, H., eds) Vol. 2, pp. 255–527, Academic Press, London
12. Calvin, N. M., and Hanawalt, P. C. (1988) *J. Bacteriol.* **170,** 2796–2801
13. Yannisch-Perron, C., Viera, J., and Messing, J. (1985) *Gene* (*Amst.*) **33,** 103–107
14. Kozak, M. (1989) *Mol. Cell. Biol.* **9,** 5134–5142
15. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989) *Gene* (*Amst.*) **77,** 61–68
16. Nilsson, I., Whitley, P., and von Heijne, G. (1994) *J. Cell Biol.* **126,** 1127–1132
17. van Geest, M., Nilsson, I., von Heijne, G., and Lolkema, J. S. (1999) *J. Biol. Chem.* **274,** 2816–2823
18. Melancon, P., and Garoff, H. (1986) *EMBO J.* **5,** 1551–1560
19. Nilsson, I., and von Heijne, G. (1992) *FEBS Lett.* **299,** 243–246
20. Tam, J. P. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 5409–5413
21. Napier, R. M., Trueman, S., Henderson, J., Boyce, J. M., Hawes, C., Fricker, M. D., and Venis, M. A. (1995) *J. Exp. Bot.* **46,** 1603–1613
22. Paretssoler, A., Pardo, J. M., and Serrano, R. (1990) *Plant Physiol.* **93,** 1654–1658
23. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
24. Sandelius, A. S., and D. J., M. (1990) *Plant Plasma Membrane: Structure, Function and Molecular Biology*, pp. 52–57, Springer-Verlag, Berlin
25. Larsson, C., Kjellbom, P., Widell, S., and Lundborg, T. (1984) *FEBS Lett.* **171,** 271–276
26. Serrano, R. (1989) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40,** 61–94
27. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395
28. Sonnhammer, E. L. L., von Heijne, G., and Krogh, A. (1998) in *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology* (Glasgow, J., Lathorp, R., Littlejohn, T., and Major, F., eds), pp. 175–182, AAAI Press, Menlo Park, CA
29. Claros, M. G., and von Heijne, G. (1994) *Comp. Appl. Biosci.* **10,** 685–686
30. Wallin, E., and von Heijne, G. (1998) *Protein Sci.* **7,** 1029–1038
31. Kaplan, H. A., Welply, J. K., and Lennarz, W. J. (1987) *Biochim. Biophys. Acta* **906,** 161–173
32. Maddy, A. H. (1976) *J. Theor. Biol.* **62,** 315–326
33. Nilsson, I.-M., and von Heijne, G. (1993) *J. Biol. Chem.* **268,** 5798–5801
34. Popov, M., Tam, L. Y., Li, J., and Reithmeier, R. A. F. (1997) *J. Biol. Chem.* **272,** 18325–18332
35. Bilgin, N., Lee, J. I., Zhu, H. Y., Dalbey, R., and von Heijne, G. (1990) *EMBO J.* **9,** 2717–2722
36. Gafvelin, G., Sakaguchi, M., Andersson, H., and von Heijne, G. (1997) *J. Biol. Chem.* **272,** 6119–6127
37. Fujiki, Y., Hubbard, A. L., Fowler, S., and Lazarow, P. B. (1982) *J. Cell Biol.* **93,** 97–102
38. von Heijne, G. (1998) *Acta Physiol. Scand.* **163,** 17–19
39. Baldwin, J. M. (1993) *EMBO J.* **12,** 1693–1703
40. van Neuren, A. S., Müller, G., Klebe, G., and Moroder, L. (1999) *J. Rec. Sig. Transduct. Res.* **19,** 341–353

41. Blobel, G. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 1496–1500
42. von Heijne, G. (1996) *Prog. Biophys. Mol. Biol.* **66,** 113–139
43. Caplan, M. J. (1991) *Curr. Top. Membr.* **39,** 37–86
44. Pryer, N. K., Wuestehube, L. J., and Schekman, R. (1992) *Annu. Rev. Biochem.* **61,** 471–516
45. Teasdale, R. D., and Jackson, M. R. (1996) *Annu. Rev. Cell Dev. Biol.* **12,** 27–54
46. von Heijne, G. (1998) *Nature* **396,** 111–112
47. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Protein Eng.* **10,** 1–6
48. Bourne, H. R. (1997) *Curr. Opin. Cell Biol.* **9,** 134–142
49. The *C. elegans* Sequencing Consortium (1998) *Science* **282,** 2012–2018
50. Horn, F., Weare, J., Beukers, M. W., Horsch, S., Bairoch, A., Chen, W., Edvardsen, O., Campagne, F., and Vriend, G. (1998) *Nucleic Acids Res.* **26,** 275–279
51. Horn, F., Bywater, R., Krause, G., Kuipers, W., Oliveira, L., Paiva, A. C. M., Sander, C., and Vriend, G. (1998) *Receptors Channels* **5,** 305–314
52. Wess, J. (1997) *FASEB J.* **11,** 346–354
53. Bockaert, J., and Pin, J. P. (1999) *EMBO J.* **18,** 1723–1729
54. Josefsson, L. G., and Rask, L. (1997) *Eur. J. Biochem.* **249,** 415–420
55. Plakidou-Dymock, S., Dymock, D., and Hooley, R. (1998) *Curr. Biol.* **8,** 315–324
56. Kehoe, D. M., Villand, P., and Somerville, S. (1999) *Trends Plant Sci.* **4,** 38–41
57. Bevan, M., Bancroft, I., Bent, E., Love, K., Goodman, H., Dean, C., Bergkamp, R., Dirkse, W., Van Staveren, M., Stiekema, W., Drost, L., Ridley, P., Hudson, S. A., Patel, K., Murphy, G., Piffanelli, P., Wedler, H., Wedler, E., Wambutt, R., Weitzenegger, T., Pohl, T. M., Terryn, N., Gielen, J., Villarroel, R., De Clerck, R., Van Montagu, M., Lecharny, A., Auborg, S., Gy, I., Kreis, M., Lao, N., Kavanagh, T., Hempel, S., Kotter, P., Entian, K. D., Rieger, M., Schaeffer, M., Funk, B., Mueller-Auer, S., Silvey, M., James, R., Montfort, A., Pons, A., Puigdomenech, P., Douka, A., Voukelatou, E., Milioni, D., Hatzopoulos, P., Piravandi, E., Obermaier, B., Hilbert, H., Dusterhoft, A., Moores, T., Jones, J. D. G., Eneva, T., Palme, K., Benes, V., Rechman, S., Ansorge, W., Cooke, R., Berger, C., Delseny, M., Voet, M., Volckaert, G., Mewes, H. W., Klosterman, S., Schueller, C., and Chalwatzis, N. (1998) *Nature* **391,** 485–488
58. Bevan, M., Bancroft, I., Mewes, H. W., Martienssen, R., and McCombie, R. (1999) *Bioessays* **21,** 110–120
59. Ma, H., Yanofsky, M. F., and Meyerowitz, E. M. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 3821–3825
60. Hedin, K. E., Duerson, K., and Clapham, D. E. (1993) *Cell. Signal.* **5,** 505–518