High Throughput Identification of Potential *Arabidopsis* Mitogen-activated Protein Kinases Substrates*[S]

Tanja Feilner‡§, Claus Hultschig‡, Justin Lee¶, Svenja Meyer‖, Richard G. H. Immink, Andrea Koenig‡, Alexandra Possling‡ ‡‡, Harald Seitz‡, Allan Beveridge‡§, Dierk Scheel¶, Dolores J. Cahill‡§, Hans Lehrach‡, Jürgen Kreutzberger‡, and Birgit Kersten‡§§**

Mitogen-activated protein kinase (MAPK) cascades are universal and highly conserved signal transduction modules in eucaryotes, including plants. These protein phosphorylation cascades link extracellular stimuli to a wide range of cellular responses. However, the underlying mechanisms are so far unknown as information about phosphorylation substrates of plant MAPKs is lacking. In this study we addressed the challenging task of identifying potential substrates for *Arabidopsis thaliana* mitogen-activated protein kinases MPK3 and MPK6, which are activated by many environmental stress factors. For this purpose, we developed a novel protein microarray-based proteomic method allowing high throughput study of protein phosphorylation. We generated protein microarrays including 1,690 *Arabidopsis* proteins, which were obtained from the expression of an almost nonredundant uniclone set derived from an inflorescence meristem cDNA expression library. Microarrays were incubated with MAPKs in the presence of radioactive ATP. Using a threshold-based quantification method to evaluate the microarray results, we were able to identify 48 potential substrates of MPK3 and 39 of MPK6. 26 of them are common for both kinases. One of the identified MPK6 substrates, 1-aminocyclopropane-1-carboxylic acid synthase-6, was just recently shown as the first plant MAPK substrate *in vivo*, demonstrating the potential of our method to identify substrates with physiological relevance. Furthermore we revealed transcription factors, transcription regulators, splicing factors, receptors, histones, and others as candidate substrates indicating that regulation in response to MAPK signaling is very complex and not restricted to the transcriptional level. Nearly all of the 48 potential MPK3 substrates were confirmed by other *in vitro* methods. As a whole, our approach makes it possible to shortlist candidate substrates of mitogen-activated protein kinases as well as those of other protein kinases for further analysis. Follow-up *in vivo* experiments are essential to evaluate their physiological relevance. ***Molecular & Cellular Proteomics 4:1558–1568, 2005.***

Upon completion of the annotation of several genomes, including the *Arabidopsis* genome (1), proteomics will greatly contribute to our understanding of gene function by systematic high throughput protein investigation (2–5). Protein microarrays have become a crucial tool in this field because they allow parallel, fast, and easy analysis of up to thousands of addressable proteins, which are printed in a systematic order with high density on coated glass slides (6–11). Analytical protein microarrays consisting of antibodies or protein antigens are increasingly used to profile proteins or antibodies, respectively, in crude protein samples of interest (6, 12). Using functional protein microarrays, proteins can be directly screened *in vitro* for a large variety of activities, including protein-protein (13, 14), protein-DNA (15, 16), protein-lipid (13), and protein-drug interactions (17, 18), under a wide range of different conditions. Furthermore previous studies have demonstrated the suitability of protein microarrays to study protein phosphorylation by kinases, however, only in a low or medium throughput manner (17, 19, 20). Zhu *et al.* (19) *e.g.* used protein microarrays bearing microwells in which 17 different protein substrates were covalently immobilized to analyze protein kinases from *Saccharomyces cerevisiae*.

In our group, we have developed the first plant (20, 21) and bacterial microarrays (22) as well as a nonredundant human protein microarray (23) for analytical and functional studies. From this technical background, we aimed, in this study, to develop a protein microarray-based method for high throughput identification of potential protein kinase substrates

From the ‡Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics (MPI-MG), Ihnestr. 73, 14195 Berlin, Germany, ¶Department of Stress and Developmental Biology, Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle, Germany, ‖RZPD German Resource Centre for Genome Research GmbH, Heubnerweg 6, 14059 Berlin, Germany, gabi.rzpd.de/, and **Business Unit Bioscience, Plant Research International, P. O. Box 16, 6700 AA Wageningen, The Netherlands

Received, January 6, 2005, and in revised form, June 1, 2005

Published, MCP Papers in Press, July 11, 2005, DOI 10.1074/mcp.M500007-MCP200

and to apply this method to reveal novel substrates of *Arabidopsis* MAPKs.[1]

MAPKs are the terminal components of the "three-kinase" modules of MAPK cascades. These kinases are activated by phosphorylation: a MAPK kinase kinase (MAPKKK), which is a serine/threonine protein kinase, phosphorylates the subsequent dual specific MAPK kinase (MAPKK), which in turn activates the MAPK by phosphorylation of a threonine as well as tyrosine residue in the "activation loop" (24). Both residues are separated by one amino acid in this loop (Thr-*X*-Tyr) (25). Downstream of activated MAPKs, which are described as serine/threonine kinases, phosphorylation events occur and may influence the regulation of genes (26). However, the phosphorylation substrates of the activated MAPKs are, especially in plants, widely unexplored.

MAPK cascades are universal and highly conserved signal transduction modules in eucaryotes, including yeasts, animals, and plants, and mediate the intracellular transmission and amplification of extracellular stimuli, resulting in the induction of appropriate biochemical and physiological cellular responses (24, 27, 28). Activation of MAPK cascades is an important mechanism for stress adaptation by control of gene expression. In plants, MAPK signaling has been implicated in abiotic as well as biotic stress situations and is associated with various physiological, developmental, and hormonal responses (29, 30). In certain plant species, this activation is an important component in host and non-host resistance against several pathogens and exhibits strong similarity to the innate immune protection systems of mammals and *Drosophila* (24, 31). Roles for MAPK activation in triggering "early" defense gene expression have been demonstrated in *Arabidopsis* (32), parsley (33), and tobacco (25) in response to pathogens or pathogen-derived elicitors. In *Arabidopsis*, a complete signaling cascade following perception of a bacterial flagellin has been elucidated recently (32). Downstream of the flagellin receptor, a leucine-rich repeat (LRR) receptor kinase, the cascade consists of MEKK1 (MAPKKK), MKK4/MKK5 (MAPKKs), and MPK3/6 (MAPKs). Signaling via this cascade results in the up-regulation of WRKY22/WRKY29 transcription factor gene expression. Despite this effort, nothing is known about the phosphorylation events and their influence on gene expression downstream of activated MAPKs. Therefore, in this study we addressed the challenging task of identifying the protein substrates of two activated *Arabidopsis* MAPKs (MPK3 and MPK6) using a novel protein microarray technique as a powerful high throughput platform.

[1] The abbreviations used are: MAPK, mitogen-activated protein kinase; ACS, 1-aminocyclopropane-1-carboxylic acid synthase; AGI, *Arabidopsis* Genome Initiative; ATM1, *A. thaliana* meristem 1; CBB, Coomassie Brilliant Blue; Cy3, indocarbocyanine; FAST, fluorescence array surface technology; LRR, leucine-rich repeat; LR reaction, recombination reaction in the GATEWAY cloning system (Invitrogen) to create expression clones, recombination of attL and attR sites; MAPKK, mitogen-activated protein kinase kinase; MAPKKK, mitogen-activated protein kinase kinase kinase; MBP, myelin basic protein; Ni-NTA, nickel nitriloacetic acid; PKA, protein kinase A; ERK, extracellular signal-regulated kinase; MEK, mitogen-activated protein kinase/extracellular signal-regulated kinase kinase; MKK, MAPK kinase; MEKK, MAPK/ERK kinase kinase.

EXPERIMENTAL PROCEDURES

Plant Material and cDNA Library Generation—*Arabidopsis thaliana* plants (ecotype Columbia-0 (Col0)) were grown under standard greenhouse conditions (22 °C and 16 h of light). Shortly after bolting, tops of young inflorescences containing the inflorescence meristems, floral meristems, and young closed floral buds were harvested. Total RNA was isolated from this material as has been described by Verwoerd *et al.* (34) followed by mRNA purification with oligo(dT)-cellulose columns according to the manufacturer's instructions (Amersham Biosciences).

The SuperScript™ plasmid system for cDNA synthesis and cloning (Invitrogen) was used for cDNA library construction following the manufacturer's protocol. The cDNA fragments were directionally ligated into the SalI/NotI cloning sites of the vector pQE30NASTattB (GenBank™ accession number AY386205) and transformed into competent *Escherichia coli* SCS1/pSE11 cells (35). Transformants were picked and kept in 384-well microtiter plates.

Selection and Rearray of Putative Expression Clones—To select putative expression clones, protein filters were generated from the arrayed cDNA library and screened with an anti-RGS-His_6 antibody (Qiagen, Hilden, Germany) as described in detail previously (20, 35).

Sequence Analysis and Selection of Uniclones—cDNAs were sequenced from the 5′-end by AGOWA GmbH (Berlin, Germany) using the pQE65 vector primer. Raw sequence trace files were passed through the PHRED base calling program (36, 37), and sequence bases with confidence values <15 were masked. Cloning vector sequence removal was performed using pregap4 (Staden Package software, version 2001 (38)). Sequences with a minimum of one of the following characteristics were eliminated: (*a*) shorter than 100 nucleotides after vector and quality clipping and (*b*) sequences highly similar to *E. coli* DNA (e-value $> e^{-20}$).

The remaining sequences were (i) analyzed by BLASTX 2.2.6 (39) with the "MIPS *Arabidopsis* Protein Database" as a reference (Munich Information Center for Protein Sequences, Munich, Germany, mips.gsf.de/), (ii) translated to the corresponding peptide sequences in the first forward frame using Emboss Transeq (40), and (iii) clustered using d2_cluster with default parameters (41) to screen for redundancy of the clone set. Using d2_cluster software sequences with a minimal overlap of 100 bp and at least 90% sequence identity are placed in the same cluster. Before clustering, the sequences were screened for interspersed repeats and low complexity DNA sequences using CrossMatch and RepBase database for repetitive *Arabidopsis* DNA sequence elements (42). The cDNA sequences were clustered together with the complete annotated *Arabidopsis* gene sequences lacking intron sequences and 5′- and 3′-untranslated sequences.

The Emboss Transeq output was screened for clones having a stop codon within the first 70 triplet codes, and the putative reading frames and the existence of 5′-UTRs were calculated from the BLASTX reports using scripts written in Perl (www.perl.com/). Clones identified as singletons by cluster analysis and at least one representative for each cluster were accepted for the uniclone set. For most of the clusters one representative each was selected. Selection criteria were: (i) full-length clone if available, (ii) smallest 5′-UTRs, and (iii) highest similarity to the corresponding gene. More than one sequence per cluster was selected for clusters that show a discrepancy between cluster and BLAST analysis, *i.e.* sequences composing one cluster with *Arabidopsis* Genome Initiative (AGI) gene codes, but

BLAST found the highest similarity to an *Arabidopsis* gene not in the same cluster.

Recombinational Cloning in 96-well Format—cDNA inserts were transferred from GATEWAY™ entry vectors into *E. coli* destination vector pQE30NASTDV by LR reaction (attL × attR recombination reaction in the GATEWAY cloning technology; for further details about this reaction please see the Invitrogen website). pQE-30NAST-DV is a derivative of pQE-30NAST (GenBank™ accession number AY386205 (pQE-30NAST-attB)) that has been modified to a GATEWAY destination vector by Invitrogen. LR reactions were carried out in 96-well format according to the manufacturer's instructions with the exception that total volume and enzyme concentration were reduced by half. Transformation of *E. coli* SCS1/pSE111 cells with 1 μl LR reaction mixture was performed by heat shock in 96-well format using 30 μl of cell suspension/well. Two clones per transformation were screened with colony PCR using vector primers. Recombinant clones were grown in 96-well microtiter plates. After adding glycerol to an end concentration of 20% (v/v), clones were stored at −80 °C.

Protein Expression and Purification in 96-well Format—Proteins were expressed in 1-ml cultures and purified (via metal chelate affinity chromatography) after lysis in denaturing lysis buffer (100 mM NaH_2PO_4, 10 mM Tris-HCl, 6 M guanidine hydrochloride, pH 8) as described previously (20). Purified proteins were analyzed visually for purity using a 15% polyacrylamide gel. Protein concentrations were determined by Bradford assay (43).

Purification and Activation of MPK3 and MPK6—The ORFs of MPK3 (At3g45640) and MPK6 (At2g43790) were amplified from first strand cDNA with modified gene-specific primers and cloned into pGEX4T-1 as BamHI/XhoI or BamHI/NotI fragments, respectively. Recombinant proteins were expressed in *E. coli* cells and purified by glutathione-agarose chromatography according to the instructions of the supplier (Sigma-Aldrich Chemie). The purified proteins were dialyzed against two changes of water (4 °C, overnight). Myelin basic protein (MBP; Sigma-Aldrich Chemie) was used as an artificial substrate to evaluate the activity of the recombinant proteins. The inclusion of 1 mM $MnCl_2$ in addition to 5 mM $MgCl_2$ in microarray-based kinase assays (see reaction details below) was found to autoactivate the recombinant MPK3 to activity levels comparable to those extracted from elicited plant material.

Generation of Protein Microarrays—The purified His-tagged proteins were positioned on FAST™ slides (Schleicher & Schuell; Batch Numbers AOBZ001, AOBZ033, AOBZ035, and AOBZ645) alongside with positive and negative controls in an 11 × 11 spotting pattern in two identical fields at a relative humidity of 75% and at 20 °C. The protein-containing plates were cooled to 8 °C during arraying. The proteins were printed with an in-house modified and extended Genetix QArray microarrayer (Genetix LTD, New Milton, UK) equipped with an optimized 4 × 4 print head of Genetix stainless steel solid pins (X2777, tip diameter of 150 μm). Pins were inked in 20 μl of protein solution for 50 ms and stamped at one position once for 100 ms using a soft touch and a spotting distance of 1 mm. After every tenth transfer and before addressing a new inking position the solid pins were washed twice for 1 s with deionized water, washed once for 3 s with 80% technical ethanol (20% bidistilled water), and dried for 3 s with oil-free, pressured air supplied at 1 bar. A mixture of MBP (end concentration of 2,000 ng/μl; Sigma-Aldrich Chemie) and mouse anti-RGS-His_6 antibody (end dilution of 1:10; Qiagen) in deionized water were chosen as positive controls for the phosphorylation experiments as well as for the detection of the proteins on the array surface. These positive controls were arranged as guide dots six times in each block of the array in an asymmetric fashion (schematic representation of the guide dots is given in Fig. 1) to assure correct orientation of each image for analysis and to justify the grids used for quantification. All purified proteins were spotted once in each of the identical fields in total twice on each array. In addition one spot position in each block was not addressed. Further controls (deionized water, PBS, 20 pmol/μl BSA, mouse anti-RGS-His_6 antibody (Qiagen) diluted 1:10 in PBS, and rabbit anti-mouse IgG3-Cy3 (Dianova, Hamburg, Germany) diluted 1:25 in PBS) were spotted three times per field (six times on each array). After completion of the spotting run, a Genepix array list file, describing the positioning of all samples on the microarray, was generated.

Immunoscreening of Protein Microarrays—The microarrays were washed for 1 h at room temperature with TBST (TBS, 0.1% (v/v) Tween 20) and than blocked for 1 h at room temperature with 2% (w/v) BSA, TBST. Mouse anti-RGS-His_6 antibody (Qiagen) was diluted 1:2,000 in blocking solution and then applied onto the arrays for 1 h at room temperature followed by two 10-min wash steps with TBST. The microarrays were further incubated for 1 h at room temperature with the respective Cy3-labeled secondary antibody (rabbit anti-mouse IgG conjugate; Dianova), which was applied with a 1:800 dilution in blocking solution. Then three wash steps of 30 min each were performed in TBST. All antibody incubation steps were carried out using a 200-μl volume underneath a coverslip in the dark. Microarrays were dried. Fluorescence signals were detected with an Affymetrix 428 microarray scanner (MWG-Biotech AG, Ebersberg, Germany) at 532 nm.

Microarray-based Kinase Assay—Protein microarrays, which were spotted the previous day, were washed in TBST for 1 h at room temperature with vigorous shaking to remove urea from the microarrays. Microarrays were blocked for 1 h at room temperature with 2% (w/v) BSA, TBST. All kinase incubations were then performed in the presence of [γ-^{33}P]ATP (25 μCi/ml; Amersham Biosciences). Protein kinase A (PKA) reactions were carried out with 12.5 μg/ml PKA from mouse (New England Biolabs GmbH, Schwalbach, Germany) in PKA buffer (50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$) for 20 min. In the case of MAPKs the microarrays were incubated with 100 ng/μl MPK3 or 200 ng/μl MPK6, respectively, in MAPK buffer (25 mM Tris-HCl, pH 7.5, 1 mM EGTA, 1 mM DTT, 5 mM $MgCl_2$, 1 mM $MnCl_2$, 20 μM ATP) for 30 min. After incubation the microarrays were washed as follows: twice for 15 min in 2× PBST (PBS, 0.1% (v/v) Tween 20), twice for 15 min in 1× PBST, once for 30 min in 1× PBST, once for 30 min in 0.5× PBST, once for 30 min in 0.1× PBST, and once for 30 min in 0.1× PBS.

Microarrays were dried and transferred into an x-ray cassette (Hypercassette, Amersham Biosciences). The enzymatic phosphorylation of the immobilized proteins was detected on the dried and Saran-covered microarrays with BAS-SR0813 imaging screens (Fujifilm). After exposition (6–48 h, depending on the specific activity of the kinase under investigation) the screens were read out with an FLA-8,000 microarray scanner (Fujifilm) at 635 nm and a spatial resolution of 10 μm.

Evaluation of Radioactive Signals and Selection of Potential Targets—The resulting images of the screens were opened in Aida Array Metrix version 3.45 (Raytest, Straubenhardt, Germany; www.raytest.de). Subsequently the Genepix array list file, describing the positions of the samples on the array, was imported. The illustrated grid of the protein samples on the array was arranged on the images according to the positions of guide dots. The spotting pattern and the relative arrangement of identical fields were entered according to the user's manual of the software package. These settings were saved as a template and used for analyzing all images. Subsequently the automatic grid-positioning feature was used, and the intensity of all spots was determined using a spot diameter of 320 μm. The average background of block spot was determined with the "mode of non-spot" evaluation mode. According to the manual the spot diameter, which is not considered for blockwise background correction, was enlarged by 4 μm. Subsequently the duplicate correlation of all spots

was determined (Fig. 3). To avoid any artifacts resulting from the precipitation of labeled kinases, only spots deviating less than 25% from the average intensity of both spots (Fig. 3, region *inside* both *green outer lines*) were considered in the subsequent analysis. Signals (spots) deviating more than 25% from the average signal intensity were excluded from the subsequent analysis by marking them in the Aida Array Metrix software package. The reproducibility of the quantification of different microarrays was verified in Aida Array Compare version 3.53. Spots showing a deviation of the background corrected signals assigned to the same protein of more than a factor of 2 between two arrays were not included in the target identification. Finally only those spots exceeding the average signal intensity of the background by at least 10 times the deviation of the background signals were considered to identify potential targets of the kinases.

On-blot Phosphorylation Assay—Proteins were separated using 15% SDS-PAGE followed by blotting the proteins on PVDF membranes and then Coomassie Brilliant Blue (CBB) staining of the gels. On-blot phosphorylations were carried out using reaction conditions as described for the microarray-based kinase assay in an appropriate volume. As positive controls MBP in different dilution steps (1:2 dilution steps from 2,000–125 ng/μl MBP) were used. For negative controls, proteins that were not identified as potential substrates in the microarray assay were used (proteins expressed from clones 311_A05, 312_H05, and 313_K06).

Testing of Protein Solubility in 96-well Format—Protein expression and purification (via metal chelate affinity chromatography) were performed manually as described by Bussow *et al.* (44) with the exception that 50 mM Tris, pH 8.0 was substituted by 50 mM Hepes, pH 8.0 for cell lysis under native conditions. Aliquots of the following samples were collected for SDS-PAGE: (i) lysates, (ii) soluble cellular proteins (supernatant after centrifugation of the lysates), and (iii) purified protein (eluates after protein purification).

Kinase Assay in Solution Using Refolded Proteins—Purification of the potential targets for verification was performed as described above (see "Protein Expression and Purification in 96-well Format") except for a scale-up of the bacteria culture volume to 5 ml and of Ni-NTA-agarose (Qiagen) to 100 μl. After the last wash step of denaturing purification, the pH of the solution was adjusted by adding 33 μl of 1 M Tris (pH 7.5) to the ~100 μl of remaining wash buffer. Consecutive dilutions by adding 130 μl, 260 μl, and 1 ml of refolding buffer (10 mM Tris, pH 7.5, containing 1 mM PMSF) were performed within a period of 30 min. The urea concentration was thus reduced to 0.5 M in the final step. After centrifugation, the buffer was removed, 1 ml of refolding buffer was added, and the samples were left on ice for another 2 h. 10 μl of the refolded proteins still attached to the Ni-NTA beads were then used for in-solution kinase assays as described previously (33).

RESULTS

Construction and Rearray of an Ordered Arabidopsis cDNA Expression Library—As a source of recombinant *Arabidopsis* proteins, we constructed a cDNA library specific for the inflorescence meristem of *Arabidopsis* in the *E. coli* expression vector pQE30NASTattB (GenBank™ accession AY386205). This vector allows isopropyl thiogalactoside-inducible expression of RGS-His_6-tagged proteins. After transformation of *E. coli* cells with the library, recombinant clones were arrayed into 384-well microtiter plates. The library consists of 40,000 clones with an average insert size of 1.06 kb. The arrayed library was gridded onto high density filter membranes and screened with an anti-RGS-His_6 antibody for putative expression clones (4,999 clones), which afterward were rearrayed robotically into a sublibrary named ATM1.

Sequence Analysis of the Library and Generation of a Uniclone Set—To create a uniclone set from this sublibrary, all of the 4,999 clones were sequenced from the 5′-end. Sequence data were passed through several analysis steps after vector clipping. 4,398 sequences remained after elimination of low quality sequences and sequences similar to *E. coli*. All sequence data can be accessed via GenBank™ by accession numbers CK117511 to CK122014.

BLASTX analysis of the sequences revealed that 61.8% of the clone inserts are in the frame of the His tag, which is comparable to the rearrayed hEx1 library (55%) (23). Furthermore it was calculated that 39.4% of the clones from the ATM1 library contain the full-length coding sequence. The d2_cluster analysis computed 2,029 different sequence clusters of which 32% are singletons, whereas 68% of the sequences are members of 637 different clusters. 85% of these clusters had a size of less than six sequences. All sequence reads were translated to amino acid sequences in the frame of the His tag, and 1,442 clones were excluded from the uniclone set because a stop codon was detected within the first 70 triplet codes of the insert DNA. This ensured that only such clones were used that express proteins with a size of at least ~8 kDa (23). 318 sequences were eliminated because their insert was calculated to be in the wrong frame by BLASTX analysis, and another 21 sequences were sorted out due to a lack of similarity to any *Arabidopsis* gene.

Finally 2,615 sequences passed all the preceding steps. 987 sequences computed to be singletons were picked out directly for the uniclone set. From the remaining 1,628 sequences belonging to 404 different clusters, 511 additional sequences were selected (one sequence from 384 clusters, and more than one from 20 clusters). These selections resulted in a uniclone set of altogether 1,498 clones. Sequence analysis results for the whole sequence dataset are provided in Supplemental Table 1 or in Supplemental Table 2 for the uniclone set.

Extension of the Uniclone Set—The uniclone set of 1,498 clones was extended by 192 full-length cDNA clones. 96 of them were described previously (21). The other 96 clones are transcription factor expression clones, which were generated from GATEWAY entry clones by LR reaction in this study. These 192 full-length clones together with the 1,498 uniclones compose the extended uniclone set of 1,690 clones.

Phosphorylation Studies with MPK3, MPK6, and PKA Using Arabidopsis Protein Microarrays—All 1,690 clones of the extended uniclone set were expressed in parallel in 96-well format. Proteins were purified by nickel chelate affinity chromatography under denaturing conditions. To control the quality of the purification 96 randomly chosen proteins were separated using 15% SDS-PAGE followed by CBB staining. 83% of the proteins were detected with a size range of 14–50 kDa (data not shown).

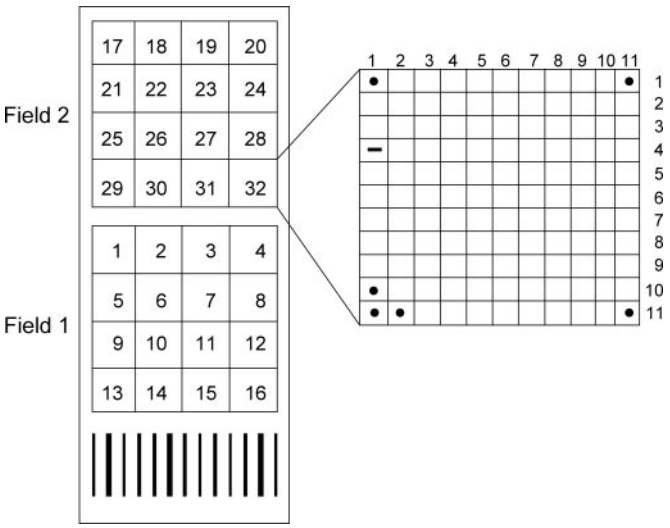

FIG. 1. **Spotting pattern of *Arabidopsis* protein microarrays.** Two identical fields were spotted per microarray. Each field consists of 16 blocks with an 11 × 11 pattern. Positions of the guide dots are marked by *filled circles*. Guide dots (mixture of MBP (2,000 ng/μl) and mouse anti-RGS-His_6 antibody (1:10)) were spotted in every block six times in an asymmetric fashion to assure correct orientation of each image for analysis and to arrange the grids used for quantification. One position per block was not addressed as indicated by a *hyphen*.

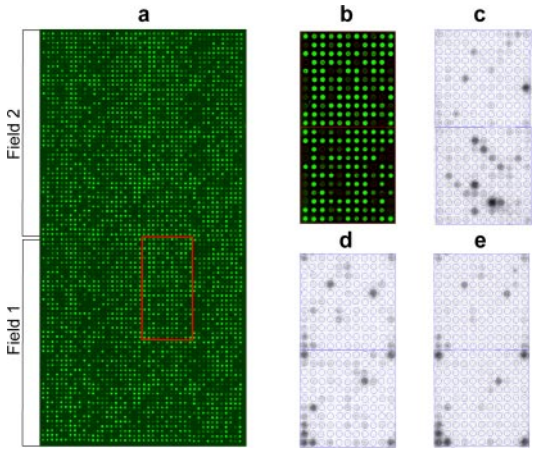

FIG. 2. **Immunoscreening (*a* and *b*) and phosphorylation studies (*c–e*) on *Arabidopsis* protein microarrays.** All proteins from the extended *Arabidopsis* uniclone set were immobilized on FAST slides and screened with an anti-RGS-His_6 antibody (*a*, whole microarray; *b*, enlargement of blocks 3 and 7 *highlighted* in *a*). Phosphorylation studies were performed on *Arabidopsis* protein microarrays with PKA (*c*), MPK3 (*d*), and MPK6 (*e*). The section of the microarray shown in *b* is given in *c–e* for the phosphorylation experiments using the same dynamic range (0–40,000) for all kinases. The PKA microarray (*c*) was exposed for 12 h, the MPK3 microarray (*d*) was exposed for 6 h, and the MPK6 microarray (*e*) was exposed for 48 h to the imaging screen.

To generate *Arabidopsis* protein microarrays, the 1,690 purified proteins of the extended uniclone set and controls were arrayed on FAST slides. The proteins were spotted once in each of the two identically fields. Each field consists of 16 blocks, and the proteins were arranged in an 11 × 11 spotting pattern in each block (see Fig. 1). The microarrays were screened with an anti-RGS-His_6 antibody (Fig. 2, *a* and *b*) to detect recombinant proteins. Furthermore the microarrays were used for the identification of phosphorylation substrates of PKA (Fig. 2*c*), MPK3 (Fig. 2*d*), or MPK6 (Fig. 2*e*), each in the presence of radioactive ATP. Mouse PKA was included as an example of a kinase from a different family to validate the specificity of our method. In addition to other controls, we used a mixture of MBP and anti-RGS-His_6 antibody as positive control and guide dot for both immunoscreening and phosphorylation studies. This control was spotted six times in each block in an asymmetric fashion as detailed in Fig. 1. This arrangement facilitates the grid adjustment for subsequent quantification. One position of every block was not addressed (Fig. 1).

For nearly all (95%) of the spotted proteins of the extended uniclone set, a signal was detected in the immunoscreen (Fig. 2*a*). The incubation with the two MAPKs resulted in signals at distinct positions including the MBP spotting positions (Fig. 2, *d* and *e*). Whereas these MAPK images are very similar, incubation with PKA resulted in a clearly different phosphorylation pattern (Fig. 2*c*). No signals were detectable in control experiments with radioactive ATP incubation without any kinase (data not shown), thus excluding that the signals observed were due to direct ATP binding or autophosphorylation of proteins with kinase domains.

To select potential substrates of *Arabidopsis* MPK3 and MPK6 based on a quantitative system, we performed two independent experiments for both kinases including two microarrays each. The recombinant MPK6 was comparatively less active, probably as a result of poor folding in the *E. coli* expression system, and hence was used at a higher concentration for the screen (see "Experimental Procedures"). Signal intensities were determined for every microarray with a high field-to-field correlation as demonstrated for one MPK3 microarray in Fig. 3. Only proteins that were identified as targets in both experiments according to our quantitative threshold-based criteria (see "Experimental Procedures") were defined as potential targets, resulting in a list of 48 or 39 for MPK3 (Table I) or MPK6 (Table II), respectively. 26 of them are common for both kinases as indicated with grey background in the tables. For comparison only, 35 substrates for mouse PKA have been identified in one experiment with two microarrays, and only a very low number of these targets overlap with the identified MAPK targets (three for MPK3, and four for MPK6) (data not shown).

To verify the potential substrates by a simple *in vitro* method, all 48 MPK3 substrates were purified, separated by SDS-PAGE, transferred onto PVDF membranes, and then incubated with MPK3 using the same assay conditions as in the microarray-based kinase assay (on-blot phosphorylation). Nearly all of the 48 potential MPK3 substrates with the exception of four samples (numbers 44, 45, 47, and 48 in Table

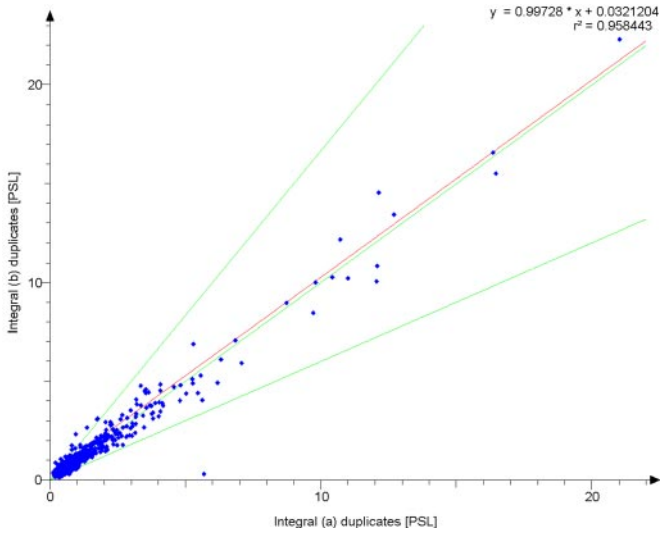

FIG. 3. **Duplicate correlation of signals of spotted proteins phosphorylated with MPK3 on one microarray.** Signal intensities of identical spots were correlated using the program Aida Array Metrix. Signals of the phosphorylated spots in field 2 (*y axis*) are plotted against the signal intensities of the corresponding proteins in field 1 (*x axis*). Each spot of the field-to-field correlation represents the signal intensity of one protein in these fields. Only duplicates deviating less than 25% from the average intensity of both spots were considered for subsequent analysis (region *inside* of both *green outer lines*). The *red line* shows the ideal distribution of theoretical 100% reproducible replicates. The regression line of experimental replicates is given in *green* (*inner line*). *PSL*, photo-stimulated luminescence; *r*, regression coefficient.

l) were phosphorylated at the membrane, whereas the negative controls were not detectable (data not shown). To further verify the targets in an independent *in vitro* assay using "native" proteins we first checked the solubility of all MPK3 and MPK6 substrate candidates in 96-well format. The tested proteins were not or only poorly soluble after expression in *E. coli* and lysis with lysozyme and therefore could not be purified under native conditions in our experiment (data not shown). We next attempted to purify the proteins, refold them while still attached via the polyhistidine tag to Ni-NTA beads, and subject them to *in vitro* kinase assays in solution. This approach is analogous to immunoprecipitation-coupled kinase assays that we routinely perform (33, 45) except that, in this case, the substrate and not the kinase is attached to the beads. As seen in Fig. 4, almost all samples with the exception of six (numbers 15, 27, 28, 44, 47, and 48 in Table I) show enhanced phosphorylation in the presence of MPK3.

DISCUSSION

Here a new proteomic method allowing high throughput identification of potential phosphorylation substrates of protein kinases has been developed using protein microarray technology. The application of this method for analyzing two *Arabidopsis* mitogen-activated protein kinases (MPK3 and MPK6) led to the selection of several novel substrate candidates of a set of ~1,700 purified *Arabidopsis* proteins. The results deliver new insights into MAPK downstream signaling.

The general feasibility of protein microarrays to detect protein phosphorylation by kinases has been demonstrated previously using different surfaces for protein immobilization: BSA-*N*-hydroxysuccinimide monolayers to which substrates were covalently attached (17), protein microarrays bearing polydimethylsiloxane-coated microwells (19), or FAST slides, which are covered with a nitrocellulose-derived polymer for non-covalent protein attachment (20). However, in these previous studies, substrate identification was carried out only in a low/medium throughput manner with regard to the number of proteins analyzed for phosphorylation on one microarray (17, 19, 20) and for most instances without quantification of the microarray results (17, 20). Therefore, in this study we aimed to develop a high throughput method. Furthermore the method should allow the identification of substrates based on quantitative criteria. The combination of an 11 × 11 spotting pattern with a microarray design using duplicates of every protein on one array (Fig. 1) allows us to immobilize theoretically 1,936 different samples in two identical fields. With this we have developed a high density protein microarray-based kinase assay. Compared with other currently available microarray methods (17, 19, 20) our assay provides the highest number of proteins that may be analyzed for phosphorylation in parallel. The efficiency of protein expression and purification methods as well as of the technology used for protein transfer to the microarray surface has been demonstrated by identifying nearly all of the recombinant proteins after screening the microarrays with an anti-RGS-His_6 antibody (Fig. 2*a*).

To detect the phosphorylation of immobilized proteins radioactively labeled ATP has been used. Initial studies were performed in which fluorescent dyes were applied for the identification of phosphorylated amino acids (46). Nevertheless the radioactive-based detection is still the most sensitive and robust detection method. Using phosphorimaging we were able to yield radioactive signals in the 11 × 11 pattern with sufficient resolution and a high field-to-field reproducibility (Fig. 3).

To manage substrate identification with a high significance, we performed two independent phosphorylation experiments for every kinase with two microarrays each. Only proteins that were detected in both experiments, taking into account the threshold-based quantitative criteria, were defined as potential targets thus reducing the number of false-positive results. We identified 48 potential substrates for MPK3 (Table I) and 39 for MPK6 (Table II). As expected, a large number of them (26 substrates) are common for both kinases. The lower number of MPK6 compared with MPK3 substrates probably reflects a lower representation of the MPK6 substrates in our set or a lower specific activity of the recombinant MPK6 used in the screen. To improve the significance of the method it would be recommendable for future studies to normalize the radio-

TABLE I

Potential MPK3 substrates identified in this study

Proteins with gray background have been identified as substrates for both MPK3 and MPK6. For substrates, which are marked in bold, or for their mammalian homologues, the involvement in signaling pathways has been described as reviewed by Yang *et al*. (28) and detailed in the text for some of them. Detailed clone information is available in supplemental Table 2. Nr., consecutive clone numbers; MIPS, Munich Information Centre for Protein Sequences; TIGR, The Institute for Genomic Research.

Nr.	Uniclone name	BLAST-AGI	Putative gene function (MIPS)	Putative gene function (TIGR 5)
1	311_B05	At5g23200	unknown protein	expressed protein
2	311_B06	At2g18020	60S ribosomal protein L2	60S ribosomal protein L8 (RPL8A)
3	**311_F11**	**At5g58620**	**putative protein**	**zinc finger (CCCH-type) family protein**
4	311_F16	At5g67360	cucumisin-like serine protease (gb|AAC18851.1)	cucumisin-like serine protease (ARA12) Asp48
5	311_H10	At2g40510	40S ribosomal protein S26	40S ribosomal protein S26 (RPS26A)
6	**312_A24**	**At5g02560**	**unknown protein**	**histone H2A, putative**
7	312_B03	At1g26740	unknown protein	expressed protein
8	312_B13	At1g77450	GRAB1-like protein	no apical meristem (NAM) family protein
9	312_D20	At1g02840	ribonucleoprotein SF-2 like protein	pre-mRNA splicing factor SF2 (SF2) / SR1 protein
10	312_F05	At3g11510	putative 40S ribosomal protein s14	40S ribosomal protein S14 (RPS14B)
11	**312_G12**	**At5g19290**	**phospholipase - like protein**	**esterase/lipase/thioesterase family protein**
12	312_G21	At2g19730	putative ribosomal protein L28	60S ribosomal protein L28 (RPL28A)
13	**312_J15**	**At4g03260**	**putative protein phosphatase regulatory subunit**	**leucine-rich repeat family protein**
14	312_P16	At4g39880	unknown protein	ribosomal protein L23 family protein
15	313_B11	At4g15000	ribosomal protein	60S ribosomal protein L27 (RPL27C)
16	**313_E14**	**At5g48990**	**unknown protein**	**kelch repeat-containing F-box family protein**
17	**313_F05**	**At5g62070**	**unknown protein**	**calmodulin-binding family protein**
18	313_F12	At3g48930	cytosolic ribosomal protein S11	40S ribosomal protein S11 (RPS11A)
19	313_G02	At1g64370	unknown protein	expressed protein
20	313_G18	At5g17870	plastid-specific ribosomal protein 6 precursor (Psrp-6) - like	plastid-specific ribosomal protein-related
21	313_H08	At2g39460	60S ribosomal protein L23A	60S ribosomal protein L23A (RPL23aA)
22	**313_J07**	**At1g52740**	**putative histone H2A**	**histone H2A, putative**
23	313_K04	At3g52580	putative ribosomal protein S14	40S ribosomal protein S14 (RPS14C)
24	313_K23	At5g48760	60S ribosomal protein L13a	60S ribosomal protein L13A (RPL13aD)
25	313_L13	At3g06730	thioredoxin, putative	thioredoxin family protein
26	313_N14	At4g39200	ribosomal protein S25	40S ribosomal protein S25 (RPS25E)
27	**313_O05**	**At1g65480**	**flowering time locus T (FT)**	**flowering locus T protein (FT)**
28	**314_E05**	**At5g65360**	**histone H3 (sp|P05203)**	**histone H3**
29	**314_F02**	**At5g65360**	**histone H3 (sp|P05203)**	**histone H3**
30	314_G05	At5g45775	ribosomal protein L11-like	60S ribosomal protein L11 (RPL11D)
31	314_G08	At5g10360	40S ribsomal protein S6	40S ribosomal protein S6 (RPS6B)
32	**314_G10**	**At5g44100**	**casein kinase I**	**casein kinase , putative**
33	314_G18	At3g04400	60S ribosomal protein L17	60S ribosomal protein L23 (RPL23C)
34	314_H17	At3g07350	unknown protein	expressed protein
35	314_I18	At1g23860	9G8-like splicing factor / SRZ-21	splicing factor RSZp21 (RSZP21) / 9G8-like SR protein (SRZ21)
36	314_J03	At1g07350	transformer-SR ribonucleoprotein, putative	transformer serine/arginine-rich ribonucleoprotein, putative
37	314_L17	At3g04400	60S ribosomal protein L17	60S ribosomal protein L23 (RPL23C)
38	314_L19	At3g04400	60S ribosomal protein L17	60S ribosomal protein L23 (RPL23C)
39	**314_N09**	**At5g65360**	**histone H3 (sp|P05203)**	**histone H3**
40	314_O11	At1g02070	unknown protein	expressed protein
41	314_P19	At4g17390	60S ribosomal protein L15 homolog	60S ribosomal protein L15 (RPL15B)
42	315_A04	At3g58700	Ribosomal protein L16	60S ribosomal protein L11 (RPL11B)
43	315_A05	At5g14320	30S ribosomal protein S13	30S ribosomal protein S13, chloroplast (CS13)
44	315_C07	At2g18020	60S ribosomal protein L2	60S ribosomal protein L8 (RPL8A)
45	315_E08	At1g03680	thioredoxin m1 (m-type precursor)	thioredoxin M-type 1, chloroplast (TRX-M1)
46	**315_F05**	**At2g02820**	**AtMYB 88**	**myb family transcription factor (MYB88)**
47	**315_G09**	**At5g66940**	**DNA binding protein-like**	**Dof-type zinc finger domain-containing protein**
48	315_I01	At1g16700	NADH:ubiquinone oxidoreductase (complex I)	NADH:ubiquinone oxidoreductase 23 kDa subunit, mitochondrial, putative

active signals with respect to the concentrations of the different transferred proteins.

We demonstrated the suitability of our quantitative criteria for substrate selection exemplarily for one kinase (MPK3) by verification of nearly all MPK3 substrates using two different *in vitro* methods. 92% of the targets were confirmed with on-blot phosphorylation (data not shown), and 88% of the substrate candidates were verified by independent *in vitro* phosphorylation of refolded proteins in solution (Fig. 4). The failure to confirm some of the potential substrates may be due to low amounts of purified proteins or because of the loss of a potential linear phosphorylation site by refolding of the protein in the assay in solution.

The high specificity of the method has been proven comparing the phosphorylation pattern of kinases from the same as well as from different families (Fig. 2). As expected, kinases from different families yield clearly different patterns, resulting in a low number of common substrates that we identified for

TABLE II
Potential MPK6 substrates identified in this study

See legend of Table I.

Nr.	Uniclone name	BLAST-AGI	Putative gene function (MIPS)	Putative gene function (TIGR 5)
1	311_B06	At2g18020	60S ribosomal protein L2	60S ribosomal protein L8 (RPL8A)
2	311_F11	At5g58620	putative protein	zinc finger (CCCH-type) family protein
3	311_F16	At5g67360	cucumisin-like serine protease (gb|AAC18851.1)	cucumisin-like serine protease (ARA12)
4	311_H07	At5g47570	unknown protein	expressed protein
5	311_H10	At2g40510	40S ribosomal protein S26	40S ribosomal protein S26 (RPS26A)
6	311_P20	At3g07110	putative 60S ribosomal protein L13A	60S ribosomal protein L13A (RPL13aA)
7	312_B13	At1g77450	GRAB1-like protein	no apical meristem (NAM) family protein
8	312_D20	At1g02840	ribonucleoprotein SF-2 like protein	pre-mRNA splicing factor SF2 (SF2) / SR1 protein
9	312_F05	At3g11510	putative 40S ribosomal protein s14	40S ribosomal protein S14 (RPS14B)
10	312_G01	At1g56220	unknown protein	dormancy/auxin associated family protein
11	312_G12	At5g19290	phospholipase - like protein	esterase/lipase/thioesterase family protein
12	312_I20	At1g22160	unknown protein	senescence-associated protein-related
13	312_P16	At4g39880	unknown protein	ribosomal protein L23 family protein
14	313_B05	At2g46020	putative SNF2 subfamily transcriptional activator	transcription regulatory protein SNF2
15	313_E14	At5g48990	unknown protein	kelch repeat-containing F-box family protein
16	313_G13	At4g11280	ACC synthase (AtACS-6)	1-aminocyclopropane-1-carboxylate synthase 6 / ACC synthase 6 (ACS6)
17	313_H08	At2g39460	60S ribosomal protein L23A	60S ribosomal protein L23A (RPL23aA) identical to GB:AF034694
18	313_K04	At3g52580	putative ribosomal protein S14	40S ribosomal protein S14 (RPS14C)
19	313_K23	At5g48760	60S ribosomal protein L13a	60S ribosomal protein L13A (RPL13aD)
20	313_L02	At3g22845	unknown protein	emp24/gp25L/p24 protein-related
21	313_L13	At3g06730	thioredoxin, putative	thioredoxin family protein
22	313_N06	At5g54630	unknown protein	zinc finger protein-related
23	313_O05	At1g65480	flowering time locus T (FT)	flowering locus T protein (FT)
24	313_P22	At5g06140	sorting nexin-like protein	phox (PX) domain-containing protein
25	314_F02	At5g65360	histone H3 (sp|P05203)	histone H3
26	314_G05	At5g45775	ribosomal protein L11-like	60 ribosomal protein L11 (RPL11D)
27	314_G10	At5g44100	casein kinase I	casein kinase I, putative
28	314_H17	At3g07350	unknown protein	expressed protein
29	314_I18	At1g23860	9G8-like splicing factor / SRZ-21	splicing factor RSZp21 (RSZP21) / 9G8-like SR protein (SRZ21)
30	314_L23	At5g21100	ascorbate oxidase-like protein	L-ascorbate oxidase, putative
31	314_N09	At5g65360	histone H3 (sp|P05203)	histone H3
32	314_N21	At3g60390	homeobox-leucine zipper protein HAT3	homeobox-leucine zipper protein 3 (HAT3) / HD-ZIP
33	314_P19	At4g17390	60S ribosomal protein L15 homolog	60S ribosomal protein L15 (RPL15B)
34	315_A04	At3g58700	Ribosomal protein L16	60S ribosomal protein L11 (RPL11B)
35	315_A05	At5g14320	30S ribosomal protein S13	30S ribosomal protein S13, chloroplast (CS13)
36	315_A10	At1g35680	clp. ribosomal large SU protein L21	50S ribosomal protein L21, chloroplast / CL21 (RPL21)
37	315_E08	At1g03680	thioredoxin m1 (m-type precursor)	thioredoxin M-type 1, chloroplast (TRX-M1)
38	315_I01	At1g16700	NADH:ubiquinone oxidoreductase (complex I)	NADH:ubiquinone oxidoreductase 23kDa subunit, mitochondrial, putative
39	315_P07	At5g43650	putative bHLH transcription factor (AtbHLH092)	basic helix-loop-helix (bHLH) family protein

these kinases (*e.g.* three common substrates for MPK3 and PKA). The relatively high number of mouse PKA substrates identified within our set of *Arabidopsis* proteins (35 substrates, data not shown) may be due to the fact that several plant protein kinases belong to the same group as PKA, the AGC group (named after PKA, PKG, and PKC; see www.nih.go.jp/mirror/Kinases/pkr/pk_catalytic/pk_hanks_seq_align_long.html), and are sharing specific sequence motifs such as the F*XX*F hydrophobic motif in the C terminus (47).

The method described here is a valuable supplementation in the spectrum of existing *in vitro* methods for substrate identification such as solid-phase phosphorylation screening of λ phage cDNA expression libraries described by Fukunaga and Hunter (48) or the method developed by Shokat and colleagues (49–51). The latter method uses kinases with modified ATP binding pockets that are accepting unnatural ATP analogues to display direct substrates of modified kinases in crude mixtures or cell lysates. In contrast to this method, no modification of the kinases is needed for our microarray-based screening. Compared with the screening of phage expression libraries, our method needs considerably smaller volumes of active kinase (200 μl for one incubation are suffi-

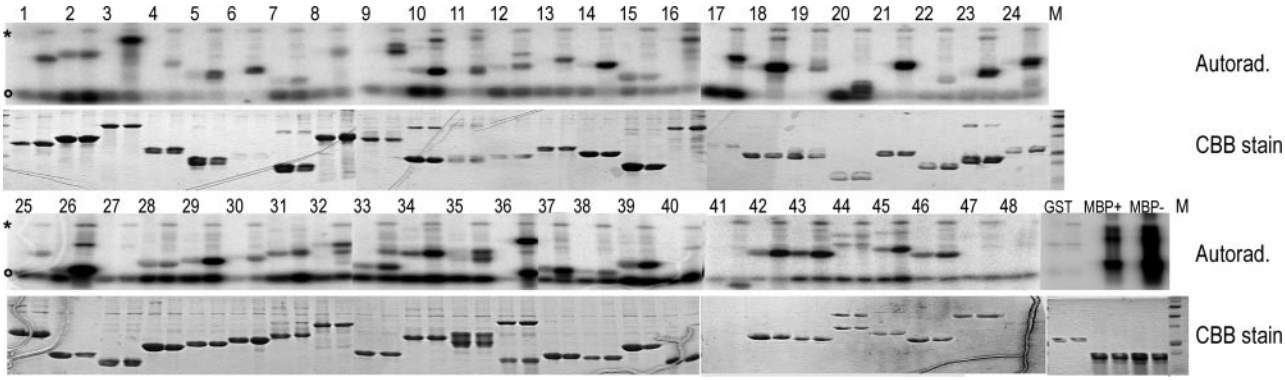

FIG. 4. **Verification of potential MPK3 targets using refolded proteins.** The 48 potential MPK3 targets were purified and refolded while still attached to Ni-NTA beads. 10 μl of the protein slurry were used for phosphorylation assays in the absence or presence (shown in *every second lane*) of active recombinant MPK3, separated by SDS-PAGE, stained with CBB, and exposed for autoradiography (*Autorad.*). The positions of the autophosphorylated MPK3 and an unspecific radioactive band across all samples are indicated on the *left* by an *asterisk* and an *empty circle*, respectively. GST and MBP were used as the respective negative and positive controls for the reactions. Possible contaminants from the Ni-NTA bead purification and refolding steps were also incubated with MBP (*MBP+*) to show the slight inhibitory effects on the assay. *M*, prestained protein markers (MBI Fermentas, St. Leon-Rot, Germany).

cient using our method; this is 1,000 times less than for the screening of phage expression libraries (52)). The screening of proteins obtained from characterized arrayed clones in our study allows direct identification of substrate candidates because a positive result is directly linked to the sequence of the respective expression clone. Further verification experiments may be performed with proteins expressed from the same clone used for the screening experiments.

In general, our assay represents a rapid *in vitro* screening for potential substrates of protein kinases. We may obtain some false-positives because we use denatured proteins as targets. In the respective native-folded protein potential phosphorylation sites are possibly not accessible for the kinase. However, the confirmation experiment with the refolded proteins points to a low number of false-positives due to this reason. Furthermore it is possible that the substrate and the kinase can interact but will never associate *in vivo*, *e.g.* because they are localized in different cellular compartments. For that reason, potential substrates identified with our method have to be verified in subsequent *in vivo* experiments; *e.g.* protein-protein interaction studies *in planta* by innovative microspectroscopic approaches such as fluorescence resonance energy transfer and fluorescence lifetime imaging microscopy (53). To gain insights into the sites of phosphorylation, antibodies against phosphorylated protein epitopes may be used for detection (54) on protein microarrays. Furthermore peptide arrays (54, 55) or MS-based methods (56, 57) may be applied in this respect.

An excellent demonstration for the suitability of our method to find substrates with biological relevance is the identification of 1-aminocyclopropane-1-carboxylic acid synthase-6 (ACS-6) as an MPK6 substrate. ACS, the rate-limiting enzyme of ethylene biosynthesis, has been described very recently as the first plant MAPK substrate *in vivo* (58). Liu *et al.* (58) found that selected isoforms of ACS including ACS-6, which we identified in our study (Table II, number 16), are substrates of MPK6. Phosphorylation of ACS led to the accumulation of the protein and to ethylene production (58). Furthermore it has been speculated that plant stress-responsive MAPKs may phosphorylate transcription factors or transcription factor-regulating proteins (24, 32) similar to their mammalian orthologs (28). Several reports support this assumption, such as the increased nuclear localization of MAPKs in parsley cells after MAPK activation (45) or the up-regulation of WRKY22/WRKY29 transcription factor expression upon activation of the flagellin MAPK cascade (32). Asai *et al.* (32) suggested that not the WRKY transcription factors themselves but a specific WRKY inhibitor may be phosphorylated and inactivated by MAPKs. In agreement with these findings, we identified several transcription factors (Table I, numbers 3, 46, and 47; Table II, numbers 2, 22, 32, and 39) as well as a transcription regulator (Table II, number 14) as potential MAPK substrates in this study. Furthermore the phosphorylation of histones (Table I, numbers 6, 22, 28, 29, and 39; Table II, numbers 25 and 31) could be involved in regulating gene transcription as it has been shown that the tail domain of histones regulates chromatin structure and hence gene transcription (59).

Several substrates, which were identified in this study, or their mammalian homologues have been described previously to be involved in signal transduction (28) as indicated in bold in Tables I and II. The identified LRR family protein (Table I, number 13) may participate as a receptor in elicitor-induced MAPK cascades due to its LRR domain (32). Flowering locus T protein (Table I, number 27; Table II, number 23) is a putative membrane-associated protein with homology to human phosphatidylethanolamine-binding protein (60). This protein is identical to Raf kinase inhibitor protein that is involved in regulation of RAF/MEK/ERK signaling pathway (61).

The calmodulin-binding family protein (Table I, number 17) may be a component of the MAPK cascade because a calm-

odulin-binding protein has been described previously as a negative regulator for stress tolerance to sodium and osmotic stress (62). A further target that has been identified for both MPKs is a casein kinase (Table I, number 32; Table II, number 27). In HeLa cells, a direct interaction of p38 MAPK and casein kinase 2 was observed, and also a stress-induced activation of casein kinase 2 by this MAPK was shown (63). All these results support the assumption that regulation in response to MAPK signaling is very complex and not restricted to the transcriptional level (28).

In conclusion, we established a powerful generic test system for the *in vitro* identification of potential protein kinase substrates by high density protein arrays followed by independent verification with refolded proteins. The application of this test system for plant MAPKs resulted in a short list of candidates for further analysis. Follow-up experiments such as *in vivo* verification and the mapping of phosphorylation sites in substrates are essential to evaluate the physiological relevance of the targets in MAPK signaling.

Acknowledgments—We thank Thomas Przewieslik and Thomas Nitsche (both MPI-MG, Berlin) for technical assistance and Dr. Ralf Stracke (Max Planck Institute for Plant Breeding Research, Cologne) for providing 96 *Arabidopsis* entry clones. We are grateful to Jasmin Bastian for reading the manuscript.

* This work was funded by the Federal Ministry of Education and Research (Grants 0312274, 0312272, and 031U102D) and the Max Planck Society. 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.

The nucleotide sequence(s) reported in the paper has been submitted to GenBank™ EBI Data Bank with accession number(s) CK117511 to CK122014.

[S] The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

§ Present address: Centre for Human Proteomics, Royal College of Surgeons in Ireland, 121 St. Stephens Green, Dublin 2, Ireland.

‡‡ Present address: Free University of Berlin, Königin-Luise-Str. 12-16, 14195 Berlin, Germany.

§§ To whom correspondence should be addressed: Dept. Neuroproteomics, Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str. 10, 13092 Berlin, Germany. Tel.: 49-30-9406-2636; Fax: 49-30-9406-2629; E-mail: b.kersten@mdc-berlin.de.

REFERENCES

1. The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* **408,** 796–815
2. Phizicky, E., Bastiaens, P. I., Zhu, H., Snyder, M., and Fields, S. (2003) Protein analysis on a proteomic scale. *Nature* **422,** 208–215
3. Kersten, B., Bürkle, L., Kuhn, E. J., Giavalisco, P., Konthur, Z., Lueking, A., Walter, G., Eickhoff, H., and Schneider, U. (2002) Large-scale plant proteomics. *Plant Mol. Biol.* **48,** 133–141
4. Kersten, B., Feilner, T., Angenendt, P., Giavalisco, P., Brenner, W., and Bürkle, L. (2004) Proteomic approaches in plant biology. *Curr. Proteomics* **1,** 131–144
5. Zhu, H., Bilgin, M., and Snyder, M. (2003) *Proteomics Annu. Rev. Biochem.* **72,** 783–812
6. Zhu, H., and Snyder, M. (2003) Protein chip technology. *Curr. Opin. Chem. Biol.* **7,** 55–63
7. Feilner, T., Kreutzberger, J., Niemann, B., Kramer, A., Possling, A., Seitz, H., and Kersten, B. (2004) Proteomic studies using microarrays. *Curr. Proteomics* **1,** 283–295
8. Labaer, J., and Ramachandran, N. (2005) Protein microarrays as tools for functional proteomics. *Curr. Opin. Chem. Biol.* **9,** 14–19
9. Poetz, O., Schwenk, J. M., Kramer, S., Stoll, D., Templin, M. F., and Joos, T. O. (2005) Protein microarrays: catching the proteome. *Mech. Ageing Dev.* **126,** 161–170
10. Glokler, J., and Angenendt, P. (2003) Protein and antibody microarray technology. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **797,** 229–240
11. Angenendt, P. (2005) Progress in protein and antibody microarray technology. *Drug Discov. Today* **10,** 503–511
12. Michaud, G. A., Salcius, M., Zhou, F., Bangham, R., Bonin, J., Guo, H., Snyder, M., Predki, P. F., and Schweitzer, B. I. (2003) Analyzing antibody specificity with whole proteome microarrays. *Nat. Biotechnol.* **21,** 1509–1512
13. Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., and Snyder, M. (2001) Global analysis of protein activities using proteome chips. *Science* **293,** 2101–2105
14. Ramachandran, N., Hainsworth, E., Bhullar, B., Eisenstein, S., Rosen, B., Lau, A. Y., Walter, J. C., and LaBaer, J. (2004) Self-assembling protein microarrays. *Science* **305,** 86–90
15. Snapyan, M., Lecocq, M., Guevel, L., Arnaud, M. C., Ghochikyan, A., and Sakanyan, V. (2003) Dissecting DNA-protein and protein-protein interactions involved in bacterial transcriptional regulation by a sensitive protein array method combining a near-infrared fluorescence detection. *Proteomics* **3,** 647–657
16. Kersten, B., Possling, A., Blaesing, F., Mirgorodskaya, E., Gobom, J., and Seitz, H. (2004) Protein microarray technology and ultraviolet crosslinking combined with mass spectrometry for the analysis of protein-DNA interactions. *Anal. Biochem.* **331,** 303–313
17. MacBeath, G., and Schreiber, S. L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science* **289,** 1760–1763
18. Zhou, F. X., Bonin, J., and Predki, P. F. (2004) Development of functional protein microarrays for drug discovery: progress and challenges. *Comb. Chem. High Throughput Screen.* **7,** 539–546
19. Zhu, H., Klemic, J. F., Chang, S., Bertone, P., Casamayor, A., Klemic, K. G., Smith, D., Gerstein, M., Reed, M. A., and Snyder, M. (2000) Analysis of yeast protein kinases using protein chips. *Nat. Genet.* **26,** 283–289
20. Kramer, A., Feilner, T., Possling, A., Radchuk, V., Weschke, W., Bürkle, L., and Kersten, B. (2004) Identification of barley CK2α targets by using the protein microarray technology. *Phytochemistry* **65,** 1777–1784
21. Kersten, B., Feilner, T., Kramer, A., Wehrmeyer, S., Possling, A., Witt, I., Zanor, M. I., Stracke, R., Luecking, A., Kreuzberger, J., Lehrach, H., and Cahill, D. J. (2003) Generation of Arabidopsis protein chips for antibody and serum screening. *Plant Mol. Biol.* **52,** 999–1010
22. Steller, S., Angenendt, P., Cahill, D. J., Heuberger, S., Lehrach, H., and Kreutzberger, J. (2005) Bacterial protein microarrays for identification of new potential diagnostic markers for Neisseria meningitidis infections. *Proteomics* **5,** 2048–2065
23. Lueking, A., Possling, A., Huber, O., Beveridge, A., Horn, M., Eickhoff, H., Schuchardt, J., Lehrach, H., and Cahill, D. J. (2003) A nonredundant human protein chip for antibody screening and serum profiling. *Mol. Cell. Proteomics* **2,** 1342–1349
24. MAPK Group (2002) Mitogen-activated protein kinase cascades in plants: a new nomenclature. *Trends Plant Sci.* **7,** 301–308
25. Yang, K. Y., Liu, Y., and Zhang, S. (2001) Activation of a mitogen-activated protein kinase pathway is involved in disease resistance in tobacco. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 741–746
26. Peck, S. C. (2003) Early phosphorylation events in biotic stress. *Curr. Opin. Plant Biol.* **6,** 334–338
27. Agrawal, G. K., Iwahashi, H., and Rakwal, R. (2003) Rice MAPKs. *Biochem. Biophys. Res. Commun.* **302,** 171–180
28. Yang, S. H., Sharrocks, A. D., and Whitmarsh, A. J. (2003) Transcriptional regulation by the MAP kinase signaling cascades. *Gene* (*Amst.*) **320,** 3–21
29. Zwerger, K., and Hirt, H. (2001) Recent advances in plant MAP kinase signalling. *Biol. Chem.* **382,** 1123–1131
30. Jonak, C., Okresz, L., Bogre, L., and Hirt, H. (2002) Complexity, cross talk and integration of plant MAP kinase signalling. *Curr. Opin. Plant Biol.* **5,**

415–424
31. Nürnberger, T., and Scheel, D. (2001) Signal transmission in the plant immune response. *Trends Plant Sci.* **6,** 372–379
32. Asai, T., Tena, G., Plotnikova, J., Willmann, M. R., Chiu, W. L., Gomez-Gomez, L., Boller, T., Ausubel, F. M., and Sheen, J. (2002) MAP kinase signalling cascade in Arabidopsis innate immunity. *Nature* **415,** 977–983
33. Kroj, T., Rudd, J. J., Nurnberger, T., Gabler, Y., Lee, J., and Scheel, D. (2003) Mitogen-activated protein kinases play an essential role in oxidative burst-independent expression of pathogenesis-related genes in parsley. *J. Biol. Chem.* **278,** 2256–2264
34. Verwoerd, T. C., Dekker, B. M., and Hoekema, A. (1989) A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Res.* **17,** 2362
35. Bussow, K., Cahill, D., Nietfeld, W., Bancroft, D., Scherzinger, E., Lehrach, H., and Walter, G. (1998) A method for global protein expression and antibody screening on high-density filters of an arrayed cDNA library. *Nucleic Acids Res.* **26,** 5007–5008
36. Ewing, B., and Green, P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8,** 186–194
37. Ewing, B., Hillier, L., Wendl, M. C., and Green, P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8,** 175–185
38. Staden, R., Beal, K. F., and Bonfield, J. K. (2000) The Staden package, 1998. *Methods Mol. Biol.* **132,** 115–130
39. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402
40. Mullan, L. J., and Bleasby, A. J. (2002) Short EMBOSS User Guide. European Molecular Biology Open Software Suite. *Brief. Bioinform.* **3,** 92–94
41. Burke, J., Davison, D., and Hide, W. (1999) d2_cluster: a validated method for clustering EST and full-length cDNA sequences. *Genome Res.* **9,** 1135–1142
42. Jurka, J. (2000) Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet.* **16,** 418–420
43. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72,** 248–254
44. Bussow, K., Quedenau, C., Sievert, V., Tischer, J., Scheich, C., Seitz, H., Hieke, B., Niesen, F. H., Gotz, F., Harttig, U., and Lehrach, H. (2004) A catalog of human cDNA expression clones and its application to structural genomics. *Genome Biol.* **5,** R71
45. Lee, J., Rudd, J. J., Macioszek, V. K., and Scheel, D. (2004) Dynamic changes in the localization of MAPK cascade components controlling pathogenesis-related (PR) gene expression during innate immunity in parsley. *J. Biol. Chem.* **279,** 22440–22448
46. Martin, K., Steinberg, T. H., Goodman, T., Schulenberg, B., Kilgore, J. A., Gee, K. R., Beechem, J. M., and Patton, W. F. (2003) Strategies and solid-phase formats for the analysis of protein and peptide phosphorylation employing a novel fluorescent phosphorylation sensor dye. *Comb. Chem. High Throughput Screen.* **6,** 331–339
47. Bogre, L., Okresz, L., Henriques, R., and Anthony, R. G. (2003) Growth signalling pathways in Arabidopsis and the AGC protein kinases. *Trends Plant Sci.* **8,** 424–431
48. Fukunaga, R., and Hunter, T. (1997) MNK1, a new MAP kinase-activated protein kinase, isolated by a novel expression screening method for identifying protein kinase substrates. *EMBO J.* **16,** 1921–1933
49. Shah, K., Liu, Y., Deirmengian, C., and Shokat, K. M. (1997) Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 3565–3570
50. Liu, Y., Shah, K., Yang, F., Witucki, L., and Shokat, K. M. (1998) Engineering Src family protein kinases with unnatural nucleotide specificity. *Chem. Biol.* **5,** 91–101
51. Eblen, S. T., Kumar, N. V., Shah, K., Henderson, M. J., Watts, C. K., Shokat, K. M., and Weber, M. J. (2003) Identification of novel ERK2 substrates through use of an engineered kinase and ATP analogs. *J. Biol. Chem.* **278,** 14926–14935
52. Fukunaga, R., and Hunter, T. (2004) *Methods in Molecular Biology: MAP Kinase Signaling Protocols*, Vol. 250, pp. 211–236, Humana Press Inc., Totowa, NJ
53. Hink, M. A., Bisselin, T., and Visser, A. J. (2002) Imaging protein-protein interactions in living cells. *Plant Mol. Biol.* **50,** 871–883
54. Lesaicherre, M. L., Uttamchandani, M., Chen, G. Y., and Yao, S. Q. (2002) Antibody-based fluorescence detection of kinase activity on a peptide array. *Bioorg. Med. Chem. Lett.* **12,** 2085–2088
55. Houseman, B. T., Huh, J. H., Kron, S. J., and Mrksich, M. (2002) Peptide chips for the quantitative evaluation of protein kinase activity. *Nat. Biotechnol.* **20,** 270–274
56. Glinski, M., Romeis, T., Witte, C. P., Wienkoop, S., and Weckwerth, W. (2003) Stable isotope labeling of phosphopeptides for multiparallel kinase target analysis and identification of phosphorylation sites. *Rapid Commun. Mass Spectrom.* **17,** 1579–1584
57. Ballif, B. A., Villen, J., Beausoleil, S. A., Schwartz, D., and Gygi, S. P. (2004) Phosphoproteomic analysis of the developing mouse brain. *Mol. Cell. Proteomics* **3,** 1093–1101
58. Liu, Y., and Zhang, S. (2004) Phosphorylation of 1-aminocyclopropane-1-carboxylic acid synthase by MPK6, a stress-responsive mitogen-activated protein kinase, induces ethylene biosynthesis in Arabidopsis. *Plant Cell* **16,** 3386–3399
59. Loury, R., and Sassone-Corsi, P. (2003) Histone phosphorylation: how to proceed. *Methods* **31,** 40–48
60. Kardailsky, I., Shukla, V. K., Ahn, J. H., Dagenais, N., Christensen, S. K., Nguyen, J. T., Chory, J., Harrison, M. J., and Weigel, D. (1999) Activation tagging of the floral inducer FT. *Science* **286,** 1962–1965
61. Yeung, K., Seitz, T., Li, S., Janosch, P., McFerran, B., Kaiser, C., Fee, F., Katsanakis, K. D., Rose, D. W., Mischak, H., Sedivy, J. M., and Kolch, W. (1999) Suppression of Raf-1 kinase activity and MAP kinase signalling by RKIP. *Nature* **401,** 173–177
62. Perruc, E., Charpenteau, M., Ramirez, B. C., Jauneau, A., Galaud, J. P., Ranjeva, R., and Ranty, B. (2004) A novel calmodulin-binding protein functions as a negative regulator of osmotic stress tolerance in Arabidopsis thaliana seedlings. *Plant J.* **38,** 410–420
63. Sayed, M., Kim, S. O., Salh, B. S., Issinger, O. G., and Pelech, S. L. (2000) Stress-induced activation of protein kinase CK2 by direct interaction with p38 mitogen-activated protein kinase. *J. Biol. Chem.* **275,** 16569–16573