Plant Molecular Biology **52:** 999–1010, 2003.

Generation of *Arabidopsis* protein chips for antibody and serum screening

B. Kersten[1,*], T. Feilner[1], A. Kramer[1], S. Wehrmeyer[1], A. Possling[1], I. Witt[2], M.I. Zanor[2], R. Stracke[3], A. Lueking[1], J. Kreutzberger[1], H. Lehrach[1] and D.J. Cahill[1]

[1]*Max Planck Institute for Molecular Genetics, Ihnestrasse 73, 14195 Berlin, Germany (*author for correspondence; e-mail kersten@molgen.mpg.de);* [2]*Max Planck Institute for Molecular Plant Physiology, Am Muehlenberg 1, 14476 Golm, Germany;* [3]*Max Planck Institute for Plant Breeding Research, Carl-von-Linne-Weg 10, 50829 Cologne, Germany*

Received 22 April 2003; accepted in revised form 26 May 2003

Key words: Arabidopsis thaliana, *Escherichia coli*, high-throughput, protein micro-array, proteomics, transcription factor

Abstract

Protein array technology has emerged as a new tool to enable ordered screening of proteins for expression and molecular interactions in high throughput. Besides classical solid-phase substrates, such as micro-titre plates and membrane filters, protein arrays have recently been devised with chip-sized supports. Several applications on protein chips have been described, but to our knowledge no studies using plant protein chips were published so far.

The aim of this study was to generate *Arabidopsis* protein chips and to demonstrate the feasibility of the protein chip technology for the investigation of antigen-antibody interactions. Therefore, *Arabidopsis* cDNAs encoding 95 different proteins were cloned into a GATEWAY-compatible *Escherichia coli* expression vector. RGS-His_6-tagged recombinant proteins were purified in high throughput and robotically arrayed onto glass slides coated either with a nitrocellulose based polymer (FAST slides) or polyacrylamide (PAA slides). Using an anti-RGS-His_6 antibody all proteins were detected on the chips. The detection limit was ca. 2–3.6 fmol per spot on FAST slides or 0.1–1.8 fmol per spot on PAA slides. The *Arabidopsis* protein chips were used for the characterisation of monoclonal antibodies or polyclonal sera. We were able to show that a monoclonal anti-TCP1 antibody and anti-MYB6 and anti-DOF11 sera bound specifically to their respective antigens and did not cross-react with the other 94 proteins including other DOF and MYB transcription factors on the chips. To enable screening of antibodies or other interacting molecules against thousands of *Arabidopsis* proteins in future, we generated an ordered cDNA expression library and started with high-throughput cloning of full-length cDNAs with GATEWAY technology.

Introduction

The *Arabidopsis* genome was completely sequenced in the year 2000. The annotation of predicted transcription units and their corresponding gene products on the genomic DNA will be further improved by the characterization of full-length cDNA collections from *Arabidopsis* (Seki *et al.*, 2002). These sequence information is a key for further functional studies of the expressed proteins. Therefore, high-throughput proteomic approaches play an increasing role (Pandey and Mann, 2000; Kersten *et al.*, 2002). A proteomic method which came up in the past few years is the protein array technology (for reviews, see Walter *et al.*, 2000; Büssow *et al.*, 2001; Cahill, 2001; Lueking *et al.*, 2001). For the generation of protein arrays with recombinant proteins a large number of cDNAs has to be cloned into an appropriate expression vector, alternatively by the construction of ordered cDNA expression libraries (Büssow *et al.*, 1998; Clark *et al.*, 1999) or by directional cloning of open reading frames (ORF) with gene-specific primers (Heyman *et al.*, 1999; Walhout *et al.*, 2000). The first approach is restricted because (1) the expressed proteins are not all full-length, and (2) full-length proteins may contain parts of the 5′-untranslated region which may lead to

artefacts in further studies. The latter approach depends on the availability of sequenced genome and the quality of gene annotation. The cloned genes have to be expressed and the proteins to be arrayed onto an immobilizing surface. Besides classical solid-phase substrates, such as micro-titre plates and membrane filters (Büssow *et al*., 1998; Lueking *et al.*, 1999), protein arrays have recently been developed with chip-sized supports, such as microscope slides coated with different gel surfaces (Arenkov *et al.*, 2000) or treated with different chemicals, such as aldehyde (MacBeath and Schreiber, 2000) to immobilize the proteins. Angenendt *et al.* (2002) compared different surfaces relative to their sensitivity, binding efficiency and signal-to-noise ratio with antibody chips.

Several applications on protein chips have been described (for reviews, see Büssow *et al*., 2001; Cahill, 2001; Lueking *et al*., 2001; Templin *et al.*, 2002; Zhu and Snyder, 2003) but, to our knowledge, no applications with protein chips with plant proteins have been published so far.

In this study we generated protein chips with 95 different *Arabidopsis* proteins including several DOF and MYB transcription factors. Therefore, the cDNAs were directionally cloned, mostly in full length, into an *E. coli* expression vector for IPTG-inducible expression of His_6-tagged proteins by using gene-specific primers. The expressed proteins were purified in high throughput and robotically arrayed onto glass slides coated either with a nitrocellulose-based polymer (FAST slides) or polyacrylamide (PAA slides). Here we compare the sensitivity of the protein detection on these chips and demonstrate the usefulness of the chips for antibody and serum screening.

Materials and methods

Modification of the expression vector

The *Escherichia coli* expression vector pQE-30NST (GenBank accession number AF074376; (Büssow *et al*., 1998)) derived from pQE-30 (Qiagen) vector for IPTG-inducible expression of RGS-His_6-tagged fusion proteins was modified by adding the restriction site *Asc*I to the vector sequence. Furthermore, this vector was made compatible to the GATEWAY system by introducing *att*B1 and *att*B2 recombination sites (Life Technologies, Karlsruhe, Germany). The resulting vector pQE-30NAST-*att*B (sequence submitted to GenBank) is schematically represented in Figure 1.

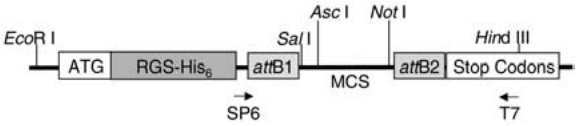

Figure 1. E. coli expression vector pQE-30NAST-*att*B (section). RGS-His_6, codons for the amino acids $RGSH_6$; *att*B1 and *att*B2, DNA recombination sequences, phage λ-based (GATEWAY); MCS, multiple cloning site.

Gene amplification

Coding sequences (mostly in full length) from *Arabidopsis* were selected from the MIPS *Arabidopsis thaliana* database (http://www.mips.biochem.mpg.de/proj/thal/). Gene-specific primers were designed with Vector NTI 6 software. Upstream of the gene-specific primer part (without start and stop codon), the primers consisted of the following nucleotide adaptor regions containing a restriction enzyme recognition site (*Asc*I in the forward primer, *Not*I in the reverse primer): 5′-AAAGGCGCGCCTG... (forward), 5′-AAAGCGGCCGCC... (reverse). The length of the gene-specific part was adjusted to an annealing temperature of ca. 55 °C. The coding sequences were amplified from different cDNA libraries or full-length cDNA clones (Table 1) with Platinium *Pfx* DNA Polymerase (Invitrogen, Groningen, Netherlands). The conditions for the gene-specific PCR were: 94 °C for 4 min; 94 °C for 30 s, 55 °C for 30 s and 72 °C for 80 s for 24 cycles; 72 °C for 3 min. PCR products were purified with a QIAquick PCR purification Kit (Qiagen, Hilden, Germany).

The following cDNA libraries from *Arabidopsis* were used as templates: the MatchMaker cDNA library (Clonetech, Heidelberg, Germany), the Minet cDNA library (Minet *et al.*, 1992), or a meristem cDNA library (Feilner *et al.*, 2002). In the case of the MatchMaker library the inserts were pre-amplified with pGAD10-vector specific primers prior to the gene-specific PCR.

Full-length cDNAs from MYB transcription factors used as templates were collected from different cDNA libraries (e.g. Labda-PRL2 cDNA library from Michigan State University; Newman *et al.*, 1994; TF cDNA collection at the Max Planck Institute for Plant Breeding Research, Cologne, unpublished). Several DOF transcription factor ORFs were PCR-amplified via gene-specific primers with integrated restriction sites (*Pme*I for the 5′ primer and *Pac*I for the 3′ primer) with first-strand cDNA pools from different *Arabidopsis* tissues. PCR products were first cloned into pUNI

Table 1a. cDNAs amplified in full length from cDNA libraries.

Lab. Code	Accession (MIPS)	CDS (bp)	Rec. Protein (AA)	Rec. Protein (kDa)	Name of the protein/predicted function
01103	At1g61520	822	321	35.3	PSI type III chlorophyll a/b-binding protein
03395	At1g19150	816	319	35.1	PSI type II chlorophyll a/b-binding protein
09137	At5g50850	1092	411	45.2	Pyruvate dehydrogenase E1 beta SU
10034	At5g17330	1509	550	60.5	Glutamate decarboxylase (GAD)
15019	At4g08900	1029	390	42.9	Arginase
15026	At2g30970	1293	478	52.6	Aspartate aminotransferase (Asp1)
15033	At5g19550	1218	453	49.8	Aspartate aminotransferase (Asp2)
15034	At5g11520	1350	497	54.7	Aspartate aminotransferase (Asp3)
18770	At2g04400	1107	416	45.8	Indole-3-glycerol phosphate synthase, Trp synth.
21214	At1g59900	1170	437	48.1	Pyruvate dehydrogenase E1 alpha SU
22964	At3g13110	1176	439	48.3	Serine acetyltransferase (Sat-1)
29699	At5g54190	1218	453	49.8	NAPH:protochlorophyllide oxidoreductase A
29785	A4g27440	1206	449	49.6	NADPH:protochlorophyllide oxidoreductase B
30298	At5g56760	939	360	39.6	Serine acetyltransferase (Sat-52)
32176	At5g64040	516	219	24.1	PSI-N
37771	At5g18170	1236	459	50.5	Glutamate dehydrognase 1 (GDH1)
43489	At4g14130	870	337	37.1	Xyloglucan endotransglycosylase-related protein (XTR7)
49937	At1g65960	1485	542	59.6	Glutamate decarboxylase (GAD2)
56635	At5g07440	1236	459	50.5	Glutamate dehydrogenase 2
95749	At1g03680	540	227	25.0	Thioredoxin m1 (m-type precursor)
176HSP	At5g12030	471	204	22.4	17.6 kDa HSP protein
30RS13	At5g14320	510	217	23.9	30S ribosomal protein S13
60STL27	At1g23290	441	194	21.3	60S ribosomal protein L27a
60SRL9	At1g33120	588	243	26.7	60S ribosomal protein L9
AAT1	At4g31990	1362	501	55.1	Aspartate aminotransferase
ADHIII	At5g43940	1140	427	47.0	Class III ADH, glutathione-dependent formaldehyde dehydrogenase
ADPRF3	n.a. /77385**	549	230	25.3	ADP-ribosylation factor 3
AMBY	A4g15210	1497	546	60.1	Beta-amylase mRNA
APT	At1g27450	552	231	25.4	Adenine phosphoribosyltransferase
APT2	At1g80050	579	240	26.4	Adenine phosphoribosyltransferase
ATPB*	n.a. /7525012**	1497	546	60.3	ATPase beta SU
BZIP	n.a. /U17887**	1110	417	45.9	bZIP protein mRNA
C24RPL2	A2g18020	777	306	33.7	60S ribosomal protein L2
CABI	At3g54890	726	289	31.8	Chlorophyll A/B-binding protein
CABM	At3g47470	756	299	32.9	Light-harvesting chlorophyll a/b binding protein
CDC2A*	At3g48750	879	340	37.4	p34(cdc2)
CKS1*	At2g27960	264	135	14.9	Cks1 protein
CPBP	n.a. /L33781**	486	209	23.0	Cab3 promoter-binding protein
CPN21	Af5g20720	762	301	33.1	Clp. CPN21-Protein (20 kDa Chaperonin)
CYCD3*	At4g34160	1104	415	45.7	Cyclin delta-3
CYSTSYN	At3g61440	1107	416	46.0	Cysteine synthase; oas5 gene
DR1B	At5g23090	480	207	22.8	Dr1
ERG9	At4g34640	1233	458	50.4	Squalene synthase
EXOPB	At3g07850	1335	492	54.1	Exopolygalacuronase (clone GBGa483)
FBA	At4g38970	1029	390	42.9	Putative fructose-biphosphate adolase
GAPB	At1g42970	1209	450	49.5	Glyceraldehyde-3-phosphate dehydrogenase B SU (GapB)
GAPA	At3g26650	1053	398	43.8	Glyceraldehyde-3-phosphate dehydrogenase A SU
GST	At2g30870	648	263	28.9	Glutathione S-transferase
ICK2*	At3g50630	630	257	28.3	Cdc2a-interacting protein
LUEAP	At2g24200	1563	568	62.5	Leucine aminopeptidase

Table 1a. Continued.

Lab. Code	Accession (MIPS)	CDS (bp)	Rec. Protein (AA)	Rec. Protein (kDa)	Name of the protein/predicted function
MTGDAS	A3g47340	1755	632	69.5	Glutamine-dependent asparagine synthetase (clone At-ASN1)
NADHUO	At5g11770	657	266	29.3	Mit. PSST SU of NADH: ubuquitone oxidoreductase (complex I)
OACYL	At4g14880	969	370	40.7	O-acetylserine lyase (At.OAS.5-8)
OEP	At5g66570	999	380	41.8	33 kDa oxygen-evolving protein
PEROX*	At3g49120	927	356	39.2	Peroxidase
PEROXp*	At3g49121	972	371	40.8	Peroxidase with putative prosequence
PHOSK	At3g12780	1443	528	58.1	Phosphoglycerate kinase, putative
PSIISBT	n.a. /X98078**	312	151	16.6	Photosystem II
RBCL*	n.a. /U91966**	1440	527	58.0	Large SU of ribulose-1,5-biphosphate carboxylase/oxygenase, clp. gene
RNABP	At2g37220	867	336	37.0	Putative RNA-binding protein
ROC4ARA	At3g62030	783	308	33.9	Nuclear encoded clp. stromal cyclophylin ROC4
RPCA	At1g43170	1170	437	48.1	Ribosomal protein
RPCL9	At3g44890	594	245	27.0	Plastid ribosomal protein CL9
RPL16	At3g58700***	549	230	25.3	Ribosomal protein L16
RPL21MR	At1g35680	663	268	29.5	clp. ribosomal large SU protein L21
RUBIACT	At2g39730	1422	521	57.5	Rubisco activase
S18RP1	At1g22780***	459	200	57.5	S18 ribosomal protein
S69727	At5g35630	1293	478	52.6	Light-regulated glutamine synthetase isoenzyme
TAGIA	At1g22070	1155	432	47.5	Transcription factor
TFIIDA	At1g55520	603	248	27.3	TFID mRNA for transcription initiation factor II (clone At-1)
TFIDB	At3g13445	603	248	27.3	TFID mRNA for transcription initiation factor II (clone At-2)
TIM	At2g21170	948	363	39.9	Putative triosephosphate isomerase
TIMC	At3g55440	759	300	33.0	Cytosolic triose phosphate isomerase
TYKYRNA	At1g16700	669	270	29.7	NADH:ubiquinone oxidoreductase (complex I)
UDGE	At1g12780	1056	399	43.9	Uridine diphosphate glucose epimerase

*amplified from the Minet library (from the Landsberg *erecta* accession) (Minet *et al*., 1992); all other genes originate from the Columbia accession.
**NCBI accession number.
***Several accession numbers exist.
n.a. not available; clp, chloroplast; mit., mitochondrial; SU, subunit; Rec, recombinant, AA, amino acid number.

(Invitrogen) and after sequence confirmation further cloned into pGreen 0229 containing the cAMV 35S promotor/terminator cassette (Hellens *et al*., 2000; http://www.pgreen.ac.uk/).

Cloning and selection of expression clones

Purified PCR products were digested with *Asc*I and *Not*I and ligated into pQE-30NAST-*att*B previously digested with the same restriction enzymes. Chemo-competent *E. coli* SCS1/pSE111 carrying lacIQ and kanamycine resistance (Büssow *et al*., 1998) were transformed with the ligation mixture by heat shock transformation. Cells were then grown for 16 h at 37 °C on LB agar plates containing ampicillin (100 μg/ml), kanamycin (25 μg/ml) and glucose (2%). Recombinant clones were selected by colony PCR with primers pQE65 (5′-TGAGCGGATAACAATTTCACACAG-3′) and pQE276 (5′-GGCAACCGAGCGTTCTGAAC-3′) prior to agarose gel electrophoresis. The selected clones were further analysed for protein overexpression by SDS-PAGE after promotor induction by IPTG. The expression clones were sequenced with primer pQE65 (DLMBC, Berlin, Germany).

Protein expression and purification

Proteins were expressed in 1 ml cultures as described (Lueking *et al*., 1999) with slight modifications. 96-well micro-titre plates with 2 ml cavities (StoreBlock, Zinser) were filled with 100 μl 2YT medium supplemented with 2% glucose, 100 μg/ml ampicillin, and 15 μg/ml kanamycin. Cultures of recombinant clones

Table 1b. cDNAs amplified in full length from single cDNA clones.

Lab. Code	Accession (MIPS)	CDS (bp)	Rec. Protein (AA)	Rec. Protein (kDa)	Name of the protein/predicted function
DOF1	At3g50410	723	288	31.7	DNA binding protein
DOF2	At2g46590	1068	403	44.3	Putative DOF zinc finger protein
DOF3	At4g21050	630	257	28.3	Prolamin box binding protein – like
DOF4	At4g21080	747	296	32.6	Transcription factor?
DOF5	At3g52440	744	295	32.5	Putative DNA-binding protein
DOF6	At4g21030	582	241	26.5	Putative protein
DOF7	At5g60200	711	284	31.2	Zinc finger protein – like
DOF8	At5g66940	675	272	29.9	DNA binding protein-like
DOF9	At5g62940	1116	419	46.1	Dof zinc finger protein – like
DOF10	At4g21040	696	279	30.7	Putative protein
DOF11	At2g34140	513	218	24.0	Putative DOF zinc finger protein
MYB6	At4g09460	711	284	31.2	DNA binding protein (clone myb6)
MYB39	At4g17780	1050	397	43.7	Putative transcription factor (MYB39)
MYB44	At5g67300	918	353	38.8	MYB-related protein
MYB77	At3g50060	915	352	38.7	Transcription factor
MYB84	At3g49690	930	357	39.3	AtMYB84 R2R3-MYB transcription factor
MYB94	At3g47600	1008	383	42.1	Putative transcription factor (MYB94)

Table 1c. cDNAs amplified as C-terminal fragments (PHYB amplified from a single cDNA clone).

Lab. Code	Accession (MIPS)	CDS (bp)	Rec. Protein (AA)	Rec. Protein (kDa)	Name of the protein/predicted function
PHOSKct	At3g12780	1200	447	49.4	Phosphoglycerate kinase
PHYBct	At2g18790	1665	602	66.2	Phytochrome B
TCP1ct	At1g22070	792	311	34.2	t-complex polypeptide 1 homologue, Top-1 homologue

were inoculated from a 96-well plate (Genetix, Christchurch, UK) that had been stored at −80 °C. After 16 h growth at 37 °C with vigorous shaking (320 rpm), 900 μl of pre-warmed medium (SB medium supplemented with 100 μg/ml ampicillin, 15 μg/ml kanamycin, 20 μg/ml thiamin) was added and the incubation was continued for 3 h. To induce protein expression, IPTG was added to a final concentration of 1 mM and incubation was continued for another 3 h. Cells were harvested by centrifugation at 1500 × g for 15 min at 4 °C and the pellets were stored at −80 °C.

Protein purification was carried out with a Qiagen BioRobot 8000 and the Ni-NTA Superflow 96 BioRobotKit (Qiagen, Hilden, Germany). Proteins were eluated in 350 μl of 10 mM Tris-HCl pH 4.5/100 mM NaH_2PO_4/8 M urea. Purified proteins were separated on a 12.5% polyacrylamide gel and protein concentrations were determined by Bradford assay (Bradford, 1976).

Generation of protein chips

Either FAST slides (Schleicher & Schuell, Dassel, Germany) or glass slides coated with a 30 μm layer of polymerized polyacrylamide/bisacrylamide (30:0.8, 8% solution) as described (Angenendt *et al.*, 2002) were used for spotting experiments. The microscope slides were placed in a Q-array system (Genetix, New Milton, UK) equipped with humidity control (55%) and 16 blunt end stainless steel pins with a tip diameter of 150 μm. The samples were spotted using a 4 × 4 or 8 × 8 spotting pattern in duplicates. Each spot was loaded once transferring a volume of 2 nl per spot.

Mouse, rat and rabbit IgG antibodies were used as controls in spotting experiments (Santa Cruz, Santa Cruz, CA).

Antibody and serum screening on protein chips

Antibody screening on FAST slides or PAA slides
The slides were blocked for 1 h at room temperature with 2% BSA (bovine serum albumin)/TBST (TBS/0.1% v/v Tween 20). Mouse anti-RGS-His_6 (Qiagen; 1:2000 dilution) or rat anti-TCP1 antibody (Affinity Bioreagents, Cholden, USA; 1:1000 dilution) was diluted 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 slides were further incubated for 1 h at room temperature with the respective Cy3-labelled secondary antibody (rabbit anti-mouse-IgG or rabbit anti-rat-IgG conjugate) (Dianova, Hamburg, Germany) which were applied with a 1:800 dilution in blocking solution. Then three wash steps of 30 min each were performed in TBST. The signal detection was performed by means of a 428 Arrayscanner System (Affimetrix, Palo Alto, CA) or a ScanArray 4000 (PerkinElmer Life Science, Köln, Germany).

All antibody incubation steps were carried out in a 200 μl volume underneath a coverslide and in the dark.

Serum screening on FAST slides
For the generation of anti-DOF11 and anti-MYB6 sera the respective RGS-His_6-tagged proteins (Table 1b) were expressed and purified under denaturating conditions on a large scale with Ni-NTA resin (Qiagen, Hilden, Germany). The purified recombinant proteins were used for the immunization of rabbits (Pineda, Berlin, Germany). Sera were collected before immunization (pre-immune sera) as well as 2 and 3 months after immunization.

The slides were blocked for 1 h at room temperature in 10% gelatin from cold water fish skin (Sigma, Saint Louis, MO)/TBST. The sera were diluted in blocking solution (1:1000 for anti-DOF11 or 1:500 for anti-MYB6 serum) and then applied onto the arrays for 1 h at room temperature. Goat anti-rabbit-IgG-Cy3 conjugate was used as a secondary antibody (Dianova). Washing was performed as described above.

Image analysis

The median spot intensity data (background substracted) were obtained with GenePixPro3.0; for further comparison the average values of the duplicated spot intensities were calculated.

Results

Cloning of Arabidopsis *genes*

For generation of His_6-tagged fusion proteins we used *E. coli* expression vector pQE-30NAST-*att*B. In addition to 23 genes coding for transcription factors (17 DOF and 6 MYB transcription factors), 117 full-length *Arabidopsis* genes were selected from the MIPS *Arabidopsis thaliana* database (http://www.mips.biochem.mpg.de/proj/thal/) randomly. Moreover, C-terminal coding sequences of the genes PHYB, TCP1 and PHOSK were selected (Table 1). With gene-specific primers for all of these 143 coding sequences, PCR amplifications were performed from cDNA libraries or single cDNA clones (Table 1). Of the 143 selected coding sequences, we managed to amplify 128 (90%) in the predicted length. Of these, 122 were cloned (95%). Of the 122 recombinant clones 95 (78%) were identified as expression clones as indicated by SDS-PAGE after over-expression in *E. coli*. In Table 1 all of the 95 expression clones are listed together with the respective proteins. By sequencing from the 5′ end we confirmed the identity of the clones.

Protein expression and purification in high-throughput

All of the 95 clones from *Arabidopsis* were expressed in parallel in a 96-well format. To complete the 96 positions in the plate MYB6 was expressed twice. Then the His_6-tagged proteins were purified robotically by nickel chelate affinity chromatography under denaturing conditions. The concentration of the purified proteins averaged 70 μg/ml corresponding to a mean yield of 24.5 μg (in 350 μl elution buffer) as determined by Bradford assay. Their predicted protein sizes are ranging from 14.9 kDa (CKS1) to 69.5 kDa (MTGDAS) (Table 1a–c). Purified proteins from 80% of the total 95 expression clones were detected by 15% SDS-PAGE followed by Coomassie staining. The results for 37 proteins are shown in Figure 2. The protein sizes deduced from the SDS gel (Figure 2) are in the range of the predicted values (Table 1a–c).

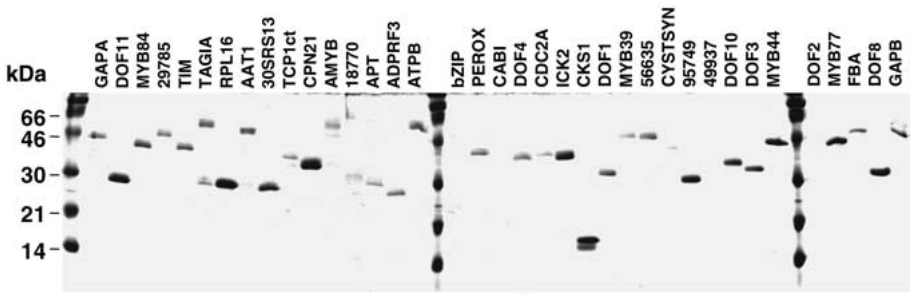

Figure 2. SDS-PAGE of 37 proteins from *Arabidopsis thaliana* (Coomassie-stained) upon high-throughput expression and purification in a 96-well formate. Approximate sizes of the protein markers are shown on the left in kDa.

Protein chip results

The 96 purified *Arabidopsis* proteins were arrayed onto FAST slides (Figure 3a) or PAA slides (data not shown) in horizontal duplicates and screened with an anti-RGS-His_6 antibody to detect His_6-tagged proteins on the chip. All recombinant proteins revealed signals over background indicating that they possess a RGS-His_6-tag. The different intensities correlate to the different concentrations of the proteins (data not shown). Solutions without protein (elution buffer, 1; PBS, 2) as well as an unrelated protein without a RGS-His_6-tag (BSA, 4) served as negative controls. These controls gave no signal in contrast to the positive control, anti-mouse-IgG-Cy3-conjugate (3). Mouse-anti-RGS-His_6 antibody (7) and mouse IgG (8) were used as controls in order to test if the secondary antibody is mouse IgG-specific and binds to the first antibody. These controls gave signals in contrast to rabbit IgG (5). The secondary anti-mouse antibody cross-reacted with rat IgG (6). To exclude unspecific binding of the secondary antibody the protein chips were incubated in blocking solution without primary anti-RGS-His_6 antibody followed by incubation with the Cy3-labeled secondary antibody. The secondary antibody did not cross-react with any of the 95 analysed *Arabidopsis* proteins (data not shown).

The protein chips with the 96 proteins and the described controls were also screened with a monoclonal anti-TCP1 antibody (Figure 3b). The image obtained with the anti-TCP1 antibody (green) was overlayed with the anti-RGS-His_6 image (red) shown in Figure 3a. As a result proteins detected with both antibodies appear yellow. This is only the case for those two spots in the main field where TCP1 was spotted indicating that the anti-TCP1 antibody specifically recognized TCP1 and did not cross-react with any of the other 94 *Arabidopsis* proteins.

The same set of protein chips (96 proteins and controls; MYB6 protein was present twice in this set) was further screened with polyclonal anti-MYB6 and anti-DOF11 sera from rabbit. In these experiments other blocking conditions than in the antibody screening experiments were used. In Figure 4 the image obtained following the anti-MYB6 serum screening (green) was overlayed with the respective anti-DOF11 image (red). Green spots were detectable exactly at the four positions where MYB6 was spotted. The red spots are located at the two DOF11 spotting positions. Four yellow spots appear in the control field. Two yellow spots arose from the positive control 3 (anti-mouse IgG-Cy3) which gave signals in both images. Rabbit IgG (control 4) also detected in both images contributed to the other two spots and neither mouse (control 6) nor rat IgG (control 8) revealed any signals indicating that the secondary antibody (anti-rabbit IgG) had bound specifically. In control experiments with pre-immune sera from the same animals which were immunized with DOF11 or MYB6, none of the 95 different *Arabidopsis* proteins was detected (data not shown).

To determine the sensitivity and detection limit, dilution series of three proteins were spotted in duplicates on FAST slides or PAA slides. The slides were screened with an anti-RGS-His_6 antibody (Figure 5a). The relative intensities of means of the duplicates are shown in Figure 5b and detection limits are compared in Table 2. The detection limit was determined to be at the lowest concentration at which the respective duplicate of the arrayed protein was clearly visible against the background. It was for all of the three proteins generally lower on PAA slides (0.1–1.8 fmol per spot) than on FAST slides (2–3.6 fmol per spot). On FAST slides, the fluorescence scaled almost linearly with the protein concentrations above 16 ng/μl (in the case of PHYB and TCP1). On PAA slides the

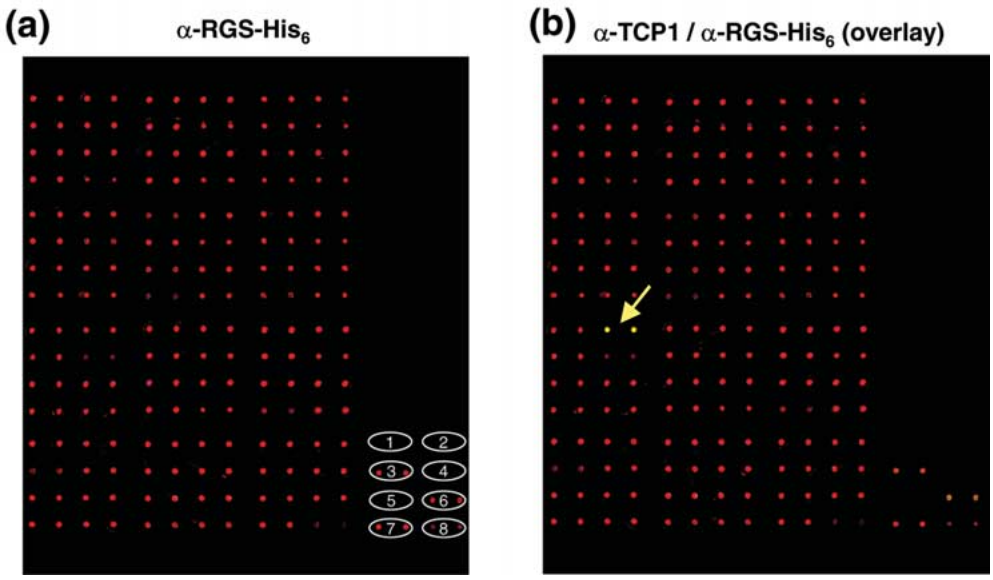

Figure 3. Protein chips with 96 *Arabidopsis* proteins screened with anti-RGS-His_6 or anti-TCP1 antibodies. (a) Anti-RGS-His_6 screened chip. (b) Overlay of the anti-RGS-His_6 (red) or anti-TCP1 (green) screening images. The purified proteins and several controls (1–8) were spotted in duplicates on FAST slides (4×4 horizontal duplicate spotting pattern, spot spacing of 1050 μm). The chips were screened with anti-RGS-His_6 (red image) or anti-TCP1 (green image) antibodies. Proteins detected with both antibodies appeared yellow in the overlayed image. Controls: 1, elution buffer; 2, PBS; 3, rabbit anti-mouse IgG-Cy3 conjugate, diluted 1:25 in PBS; 4, BSA, 20 pmol/μl in PBS; 5, normal rabbit IgG, diluted 1:10 in PBS; 6, normal rat IgG, diluted 1:10 in PBS; 7, mouse anti-RGS-His_6 antibody, diluted 1:10 in PBS; 8, normal mouse IgG, diluted 1:10 in PBS.

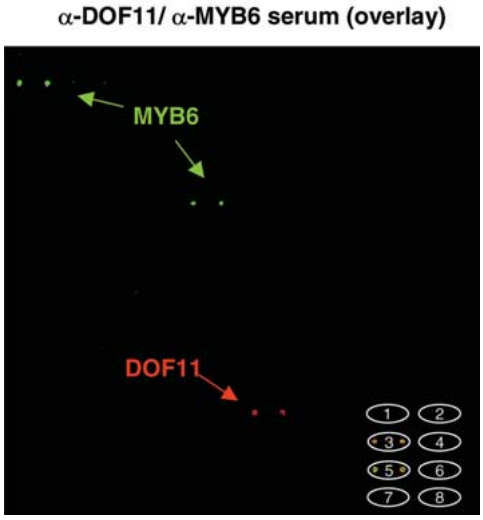

Figure 4. Overlay of two images from protein chips with 96 *Arabidopsis* proteins (MYB 6 twice in this set) and controls screened with anti-DOF11 or anti-MYB6 polyclonal sera, respectively. All samples were spotted as described in the legend to Figure 3. The chips were screened with anti-MYB6 (green image) as well as with anti-DOF11 (red image) polyclonal sera from rabbit (2 months after immunization). The figure shows an overlay of both images. Controls detected with both antibodies appear yellow in the overlayed image.

fluorescence of the spots became saturated at concentrations above 31 ng/l. The limits for the detection of PHYB and TCP1 with specific antibodies against these proteins can be derived from Figure 6. TCP1 is detectable on PAA slides above a concentration of 31 ng/μl which is equal to the detection limit for TCP1 with an anti-RGS-His_6 antibody (Table 2). The detection limit of PHYB with the specific antibody is also comparable to the limit obtained with the anti-RGS-His_6 antibody on the same surface (62 ng/μl on FAST slides).

Discussion

Protein chip technology has been applied previously to analyse antigen-antibody interactions (Arenkov *et al.*, 2000; Haab *et al.*, 2001; Robinson *et al.*, 2002) including studies from our group (Lueking *et al.*, 1999; Cahill, 2001; Angenendt *et al.*, 2002). This technology has been shown to be a useful tool to analyse enzyme-substrate reactions (Arenkov *et al.*, 2000; MacBeath and Schreiber, 2000; Zhu *et al.*, 2000). For some examples the interaction of immobilized proteins with other molecules such as proteins (MacBeath and Schreiber, 2000; Zhu and Snyder, 2001), lipids

Table 2. Detection limits of three recombinant proteins from *Arabidopsis* with anti-RGS-His_6 antibody on PAA and FAST slides.

		Detection limit			
Type of slides	Protein spotted	fmol/spot	pg/spot	pmol/μl	ng/μl
PAA	GAPA	0.7	32	0.4	16
	PHYB	0.1	8	0.1	4
	TCP1	1.8	62	0.9	31
FAST^{TM}	GAPA	2.9	124	1.4	62
	PHYB	2.0	124	1.0	62
	TCP1	3.6	124	1.8	62

(Zhu and Snyder, 2001) or small molecules (MacBeath and Schreiber, 2000) has been demonstrated with the use of protein chips.

In this study we applied protein chip technology to analyse plant proteins. We generated 95 different expression clones from *Arabidopsis thaliana* (92 of them expressing full-length cDNAs) including several MYB and DOF transcription factors. Therefore we used an *E. coli* expression vector compatible with the GATEWAY system, which allows an easy transfer of the cloned cDNAs into other expression systems in future studies. *E. coli* was used as a robust, convenient and inexpensive expression host with suitable growing properties in high-throughput format.

To generate expression clones, 143 coding sequences from *Arabidopsis* (140 full-length) were selected randomly. We managed to amplify 128 (90%) of them in the predicted length from cDNA libraries or single cDNA clones, and to also clone 122 of them (95% of the PCR products). Failure in amplification may be due to non-optimal PCR primer sequences or PCR conditions. Other reasons may be that the respective cDNA was missing or not full-length in the template library. Failure in cloning of the ORF into our *E. coli* expression vector may be due to the toxicity of the respective protein to the host.

In this study, 95 of the 122 recombinant *Arabidopsis* clones (78%) were identified as expression clones as indicated by SDS-PAGE upon over-expression in *E. coli*. This rate is an excellent expression rate compared to published results (Braun *et al.*, 2002). Braun *et al.* (2002) were able to express and purify proteins under denaturing conditions from 60% of 336 randomly selected human cDNA clones in *E. coli* using a His_6-tag. Failure of expression in our study may be due to the toxicity of the expression product to *E. coli*, to the instability of the protein, or to differences in the codon usage between *Arabidopsis* and *E. coli*. Furthermore, proteins with low expression levels were probably not detected by SDS-PAGE. A decrease in expression levels with increasing protein size has been described (Braun *et al.*, 2002).

In our study, we purified the proteins after high-throughput expression under denaturing conditions by Ni-NTA-immobilized metal affintiy chromatography performed robotically. The average yield of purified protein was 24.5 μg (70 μg/ml) and we were able to detect 80% of the 95 proteins after SDS-PAGE and Coomassie staining.

For the generation of protein chips the purified proteins were robotically arrayed onto glass slides coated either with a nitrocellulose-based polymer (FAST slides) or polyacrylamide (PAA slides). Due to their three-dimensional structure, both surfaces show a much higher protein-binding capacity than two-dimensional surfaces such as poly-L-lysin or aldehyde slides (Angenendt *et al.*, 2002). Moreover, reagents that may increase binding capacity or change the surrounding to a more physiological state, such as glycerol, can be easily added to PAA slides (Arenkov *et al.*, 2000). We observed that horseradish peroxidase (HRP) and HRP-labelled antibodies spotted onto PAA or FAST slides retained their enzymatic activity for at least two weeks when stored at 4 °C (data not shown). The functionality of three-dimensional protein chips has also been shown by Arenkov *et al.* (2000) who used a gel photo- or persulfate-induced co-polymerization technique to produce micro-arrays of gel pads for the immobilization of proteins.

In our study, all of the 95 RGS-His_6-tagged *Arabidopsis* proteins spotted were detected by an anti-RGS-His_6 antibody against a low background on

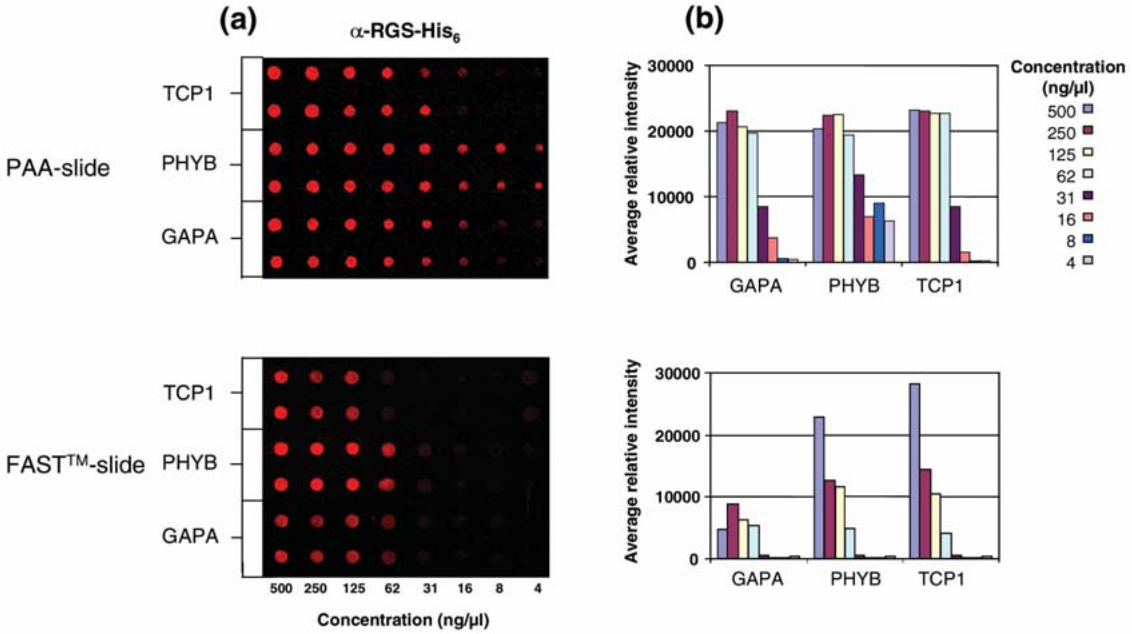

Figure 5. Sensitivity of His_6-tagged protein detection on protein chips using different surfaces (PAA/FAST slides). (a) Protein chips with three *Arabidopsis* proteins in dilution series, screened with an anti-RGS-His_6 antibody. (b) Relative intensities of means of duplicates calculated from the image shown in (a). The purified proteins were diluted in elution buffer. The same protein dilutions were spotted in duplicate (8 × 8 spotting pattern) onto PAA or FAST slides, respectively. The PAA slides were incubated following the protocol described in Materials and methods. The FAST slides were blocked with a 1× PBS/0.5% v/v Tween 20 solution for 2 h at room temperature. Anti-RGS-His_6 antibody (1:2000 dilution in blocking solution) was then applied onto the chips for 2 h at room temperature, followed by two 10 min wash steps with blocking solution. The slides were further incubated with rabbit anti-mouse IgG-Cy3 conjugate (1:800 dilution in blocking solution) for 1 h at room temperature. The slides were washed twice with 1× PBS/0.5% v/v Tween 20 for 30 min and then twice with 1× PBS for 20 min at room temperature before the signal detection was performed.

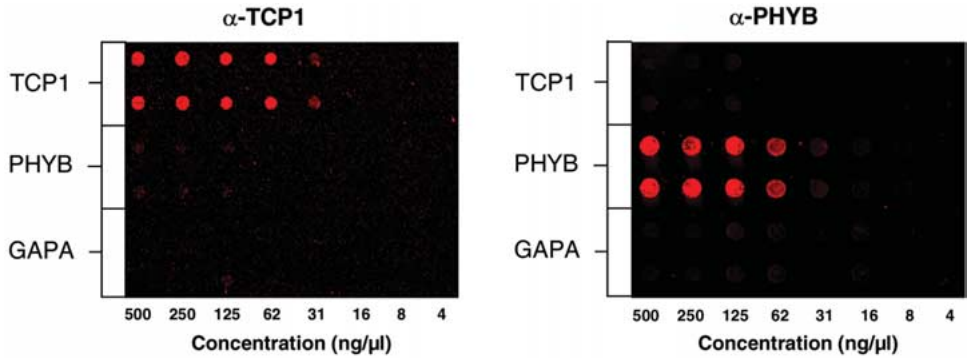

Figure 6. Detection of specific *Arabidopsis* proteins in dilution series on protein chips with monoclonal antibodies. The purified proteins were diluted with elution buffer and spotted in duplicate (8 × 8 spotting pattern) on PAA slides or FAST slides. The chips were screened with an anti-TCP1 antibody (PAA slide) as described in Materials and methods or the mouse-anti-PHYB antibody Pea-25 (Cordonnier *et al.*, 1986) (FAST slide) in a 1:1000 dilution. The procedure is the same as described for the anti-RGS-His_6 antibody in the legend to Figure 5.

FAST or PAA slides. This demonstrates that the expression and purification steps used are sufficient for detecting the proteins on these surfaces.The detection limit was 2–3.6 fmol per spot (124 pg per spot) on FAST slides or 0.1–1.8 fmol/spot (8–62 pg per spot) (Table 2) on PAA slides. On FAST slides detection limits of only 1 pg per spot have been observed on screening of *N. meningitidis* proteins with post-vaccination sera when using similar protein purification and spotting conditions (J. Kreutzberger, data not shown). This lower detection limit determined with polyclonal sera instead of a monoclonal antibody as

in our study can be explained by the fact that much more epitopes of the spotted protein are recognized by a polyclonal serum than by a monoclonal antibody.

The *Arabidopsis* protein chips with the 95 different proteins were used for characterization of the specificity and cross-reactivity of monoclonal antibodies and new polyclonal sera in a proof-of-principle study. The sera were generated against recombinant MYB6 and DOF11. We could show that a monoclonal anti-TCP1 antibody and the anti-MYB6 and anti-DOF11 sera bound specifically to their respective antigens and did not cross-react with the other 94 proteins, including other DOF and MYB transcription factors on the chips.

The protein chip technology has several advantages over other methods used for antibody or serum screening, such as western blotting. The technology enables the screening of thousands of defined proteins simultaneously. (For example, when we used a Q-array robot with a 13×13 spotting pattern, we are able to spot 2704 different proteins in one field of a chip.) Every arrayed protein is directly linked to the corresponding cDNA clone which allows the identification of positively reacting proteins very easily. Compared to western blotting, protein chip technology has the same specificity but a higher sensitivity (J. Kreutzberger, data not shown). Moreover, only a very small volume of the solutions containing the interaction partners is necessary for the incubation with protein chips (200 μl solution for one chip in this study). Another advantage of protein chip technology is that it can be easily automated.

To enable screening of antibodies, sera or other interacting molecules against thousands of *Arabidopsis* proteins in the future, we generated an ordered cDNA expression library (Feilner *et al.*, 2002) and started with high-throughput cloning of cDNAs in full length with GATEWAY technology. We are developing automated refolding procedures upon expression in *E. coli* for functional studies with *Arabidopsis* protein chips. For cases of limited or no expression in *E. coli*, we have developed two alternative expression systems in yeast (Lueking *et al.*, 2000; Holz *et al.*, 2001). We will use systems for *in vitro* protein expression directly from PCR products with the RTS 100-system (Roche Applied Science).

In this study, we have shown that protein chip technology has the potential of becoming a very important new tool for the analysis of *Arabidopsis* proteins.

Acknowledgements

We are grateful to Stefan Langhammer for expert technical assistance. We thank Jon Hughes for introducing the *Asc*I site into the expression vector and for providing the PHYB-cDNA clone and the anti-PHYB antibody Pea 25. This work was performed in partial fulfilment of the requirements for the Ph.D. of Tanja Feilner, FU Berlin, Germany.

This work was founded by the German Ministry for Research (BMBF grant 0312274, 'Large scale automated plant proteomics') and the Max Planck Society.

References

Angenendt, P., Glökler, J., Murphy, D., Lehrach, H. and Cahill, D.J. 2002. Towards optimized antibody microarrays: a comparison of current microarray support materials. Anal. Biochem. 309: 253–260.

Arenkov, P., Kukhtin, A., Gemmell, A., Voloshchuk, S., Chupeeva, V. and Mirzabekov, A. 2000. Protein microchips: use for immunoassay and enzymatic reactions. Anal. Biochem. 278: 123–131.

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.

Braun, P., Hu, Y., Shen, B., Halleck, A., Koundinya, M., Harlow, E. and LaBaer, J. 2002. Proteome-scale purification of human proteins from bacteria. Proc. Natl. Acad. Sci. USA 99: 2654–2659.

Büssow, 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. Nucl. Acids Res. 26: 5007–5008.

Büssow, K., Konthur, Z., Lueking, A., Lehrach, H. and Walter, G. 2001. Protein array technology. Potential use in medical diagnostics. Am. J. Pharmacogenomics 1: 37–43.

Cahill, D.J. 2001. Protein and antibody arrays and their medical applications. J. Immunol. Meth. 250: 81–91.

Clark, M.D., Panopoulou, G.D., Cahill, D.J., Büssow, K. and Lehrach, H. 1999. Construction and analysis of arrayed cDNA libraries [Review]. Meth. Enzymol. 303: 205–233.

Cordonnier, M.M., Greppin, H. and Pratt, L.H. 1986. Identification of a highly conserved domain on phytochrome from angiosperms to algae. Plant Physiol. 80: 982–987.

Feilner, T., Kersten, B., Immink, R.G.H., Wehrmeyer, S., Gutjahr, C., Horn, S., Possling, A., Angenent, G.C., Lehrach, H. and Cahill, D.J. 2002. Construction of an ordered *Arabidopsis* cDNA expression library for the generation of protein arrays and chips. Abstract. 1st Plant-GEMs Meeting, 29 September–2 October 2002, Berlin, Germany.

Haab, B.B., Dunham, M.J. and Brown, P.O. 2001. Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. Genome Biol. 2: 1–13.

Hellens, R.P., Edwards, E.A., Leyland, N.R., Bean, S. and Mullineaux, P.M. 2000. pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. Plant Mol. Biol. 42: 819–832.

Heyman, J.A., Cornthwaite, J., Foncerrada, L., Gilmore, J.R., Gontang, E., Hartman, K.J., Hernandez, C.L., Hood, R., Hull, H.M., Lee, W.Y., Marcil, R., Marsh, E.J., Mudd, K.M., Patino, M.J., Purcell, T.J., Rowland, J.J., Sindici, M.L. and Hoeffler, J.P. 1999. Genome-scale cloning and expression of individual open reading frames using topoisomerase I-mediated ligation. Genome Res. 9: 383–392.

Holz, C., Lueking, A., Bovekamp, L., Gutjahr, C., Bolotina, N., Lehrach, H. and Cahill, D.J. 2001. A human cDNA expression library in yeast enriched for open reading frames. Genome Res. 11: 1730–1735.

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.

Lueking, A., Horn, M., Eickhoff, H., Büssow, K., Lehrach, H. and Walter, G. 1999. Protein microarrays for gene expression and antibody screening. Anal. Biochem. 270: 103–111.

Lueking, A., Holz, C., Gotthold, C., Lehrach, H. and Cahill, D.J. 2000. A novel system for dual protein expression in Pichia pastoris and Escherichia coli. Protein Expr. Purific. 20: 372–378.

Lueking, A., Konthur, Z., Eickhoff, H., Büssow, K., Lehrach, H. and Cahill, D.J. 2001. Protein microarrays: a tool for the post-genomic era. Curr. Genom. 2: 151–159.

MacBeath, G. and Schreiber, S.L. 2000. Printing proteins as microarrays for high-throughput function determination. Science 289: 1760–1763.

Minet, M., Dufour, M.E. and Lacroute, F. 1992. Complementation of *Saccharomyces cerevisiae* auxotrophic mutants by *Arabidopsis thaliana* cDNAs. Plant J. 2: 417–422.

Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S. and Thomashow, M. 1994. Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. Plant Physiol. 106: 1241–1255.

Pandey, A. and Mann, M. 2000. Proteomics to study genes and genomes. Nature 405: 837–846.

Robinson, W.H., DiGennaro, C., Hueber, W., Haab, B.B., Kamachi, M., Dean, E.J., Fournel, S., Fong, D., Genovese, M.C., de Vegvar, H.E., Skriner, K., Hirschberg, D.L., Morris, R.I., Muller, S., Pruijn, G.J., van Venrooij, W.J., Smolen, J.S., Brown, P.O., Steinman, L. and Utz, P.J. 2002. Autoantigen microarrays for multiplex characterization of autoantibody responses. Nature Med. 8: 295–301.

Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M., Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K. 2002. Functional annotation of a full-length *Arabidopsis* cDNA collection. Science 296: 141–145.

Templin, M.F., Stoll, D., Schrenk, M., Traub, P.C., Vohringer, C.F. and Joos, T.O. 2002. Protein microarray technology. Trends Biotechnol. 20: 160–166.

Walhout, A.J., Temple, G.F., Brasch, M.A., Hartley, J.L., Lorson, M.A., van den Heuvel, S. and Vidal, M. 2000. GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. Meth. Enzymol. 328: 575–592.

Walter, G., Büssow, K., Cahill, D.J., Lueking, A. and Lehrach, H. 2000. Protein arrays for gene expression and molecular interaction screening. Curr. Opin. Microbiol. 3: 298–302.

Zhu, H. and Snyder, M. 2001. Protein arrays and microarrays. Curr. Opin. Chem. Biol. 5: 40–45.

Zhu, H. and Snyder, M. 2003. Protein chip technology. Curr. Opin. Chem. Biol. 7: 55–63.

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. Nature Genet. 26: 283–289.