2

MATERIAL AND METHODS

Methods not performed with commercially available kits are mostly based on standard protocols taken from “Molecular Cloning: A Laboratory Manual” (Sambrook *et al.* 1989). However, as some protocols have been more or less modified, they are described in more detail in this chapter.

If not stated otherwise, chemicals were purchased from Sigma-Aldrich Chemie GmbH (Taufkirchen, Germany).

2.1 Methods in molecular biology

2.1.1 Culture of *E. coli* cells

Culture of *E. coli* bacteria was performed in liquid LB medium (Luria-Bertani medium, consisting of 10 g/l Bacto Tryptone, 5 g/l Bacto Yeast Extract and 10 g/l NaCl at pH 7, obtained as powder from Q-BioGene, dissolved and sterilized by autoclaving). For solid LB-agar plates, 1.5% agarose (Q-BioGene) was added before autoclaving. Bacterial cultures for preparation of low-copy plasmid DNA were grown in 2YT rich medium (2x yeast tryptone medium, consisting of 16 g/l Bacto Tryptone, 10 g/l Bacto Yeast Extract and 5 g/l NaCl at pH 7, obtained as powder from Q-BioGene, dissolved and sterilized by autoclaving).

Antibiotics for growth selection were used in the following concentrations: Ampicillin 100 µg/ml, Kanamycin 50 µg/ml, Chloramphenicol 12.5 µg/ml and Zeocin 50 µg/ml. For Zeocin, low-salt LB medium was used containing only 5 g/l NaCl. Bacteria were grown at 37°C with 220 rpm shaking until the optical density at 595 nm reached the desired value.

For long-term storage of *E. coli* cells, glycerol stocks were prepared by adding sterile 45% (v/v) glycerol (Merck) to 10% final concentration, freezing on dry ice and storing the glycerol stocks at -80°C.

2.1.2 Transformation of *E. coli*

Heat-shock transformation-competent *E. coli* cells were prepared using a high-competence protocol (Inoue *et al.* 1990). 100 ml of SOB medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$) were inoculated with a single *E. coli* colony in a 500 ml flask. Cells were grown at 30°C with vigorous shaking. When the OD_{600} reached 0.6, the flask was placed on ice for 10 min in a pre-cooled 250 ml centrifuge bottle. Cells were pelleted with 2,000 g for 8 minutes. The supernatant was removed and the cells gently resuspended in 30 ml of ice-cold TB (10 mM Pipes, 15 mM $CaCl_2$, 250 mM KCl, pH 6.7 with KOH, then $MnCl_2$ to 55 mM final concentration, filtered sterile). After incubation on ice for 10 min, cells were pelleted with 2000 g for 8 min and resuspended gently in 8 ml TB. 650 µl DMSO was added to a final concentration of 7%. After incubation on ice for further 10 min, the suspension was aliquoted, frozen in liquid nitrogen and stored at -80°C.

Competent cells were thawed slowly on ice (~10 min). 50 µl of competent cells were added to 1-2 µl of DNA solution in 8-strips. After incubation on ice for 30 min, heat-shock was performed at 42°C (45 sec for DH5α, 120 sec for BL21 derivates). Tubes were returned to ice, then 150 µl of room-temperature SOC medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM $MgSO_4$, 20 mM glucose) was added. After incubation at 37°C for 45-60 minutes, bacteria were plated onto selective LB agar and incubated overnight at 37°C.

2.1.3 Isolation of plasmid DNA from *E. coli*

Low-throughput preparation of plasmid DNA

For small numbers of samples (up to 24 minipreps), the QIAprep Spin Miniprep Kit (Qiagen) was used according to the manufacturer's instructions. Bacterial pellets from 5 ml overnight cultures in selective LB medium were resuspended in 250 µl of buffer P1 and transferred to a microcentrifuge tube. 250 µl of buffer P2 was added and the tube gently inverted 4-6 times to mix. After addition of 350 µl of buffer N3, the tube was inverted immediately but gently 4-6 times. After centrifugation for 10 min at 10,000 g in a tabletop centrifuge, a compact white pellet formed. The supernatant from the centrifugation was applied to the QIAprep column by decanting, and the column was centrifuged for 30-60 sec at 10,000 g. The flow-through was discarded, the spin column washed by adding 500 µl of buffer PB and centrifuged for 30-60 sec

at 10,000 g. Again, the flow-through was discarded, the spin column washed by adding 750 µl of buffer PE and centrifugation for 30-60 sec at 10,000 g. After removal of residual wash buffer, the spin column was placed in a clean 1.5 ml microfuge tube. To elute DNA, 30 µl of EB (10 mM Tris-Cl, pH 8.5) was added to the center of each column, let stand for 1 min, and centrifuged for 1 min. Miniprep DNA was stored at -20°C.

High-throughput preparation of plasmid DNA

For larger numbers of samples (more than 24 minipreps), plasmid DNA preparations were performed using an in-house miniprep robot (KURABO PI-100Σ) according to the manufacturer's instructions. The following buffers were used for the miniprep protocol based on alkaline lysis and isopropanol precipitation: (1) 1xTE with 100 µg/ml RNase A for resuspension of the bacterial pellet, (2) 0.2 M NaOH with 1% SDS for alkaline lysis, (3) 1.65 M potassium acetate with 3.35 M acetic acid together with (4) 50% (v/v) ethanol with 50% (v/v) 2-propanol for DNA precipitation, (5) 70% ethanol for washing and (6) 1xTE for resuspension of plasmid DNA.

2.1.4 Agarose gel electrophoresis

Samples of nucleic acids (mainly genomic DNA and PCR products) were routinely separated according to migration behavior during agarose gel electrophoresis. Depending on the size of the fragments to be analyzed, the concentration of agarose in the gel ranged from 0.8% to 1.5%. Ethidium bromide was added to the gel in a concentration of 0.1 µg/ml for visualization. Electrophoresis was carried out in TAE buffer (40 mM Tris-acetate pH 8.2, 1 mM EDTA) using a voltage limit. Voltage and duration of the run were chosen according to the desired quality and resolution of the picture taken while exciting UV fluorescence of intercalated ethidium bromide molecules.

2.1.5 Reverse Transcription

Complementary DNA (cDNA) was generated from RNA using SuperScript II RNase H- Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions. Template RNAs were adjusted to a concentration of 1 µg/µl. In a 12 µl volume, 3 µg RNA were mixed with 7.5 µl Rnase-free water, 150 ng random hexamer primer and 10 pmol of each dNTP. The RNA was denatured at 65°C for 5 min and quickly chilled

on ice. After brief centrifugation to collect the contents, 7 µl mix for first-strand synthesis was added composed of 4 µl 5x first-strand buffer, 2 µl 0.1 M DTT and 40 U RNaseOUT Ribonuclease Inhibitor. Contents were mixed gently and distributed à 19 µl into nuclease-free 8-strips. Annealing of random primers took place at 25°C for 10 min, then a prewarming step followed at 42°C for 2 min. The mixes were transferred to new 8-strips with 1 µl (200 U) of SuperScript II for each sample and mixed by pipetting gently up and down. The RT reaction was performed at 42°C for 50 min. Reverse transcriptase was heat-inactivated at 70°C for 15 min. 1 µl (2 U) of *E. coli* RNase H was added to each sample to remove RNA complementary to the cDNA. RNA degradation was carried out at 37°C for 20 min. The resulting first-strand cDNA was frozen on dry ice and stored at -80°C.

2.1.6 Polymerase chain reaction (PCR)

PCR reactions for specific amplification of DNA sequences were carried out following three different standard protocols according to the aim of the PCR. For preparative purposes (ORF cloning), a commercial high-fidelity polymerase was used, whereas in-house produced low-fidelity Taq polymerase was used in analytical PCR reactions for economic reasons. The following PCR protocols were used:

(a) Gene-specific PCR for Gateway cloning of open reading frames

Gene-specific primer pairs with annealing temperatures of approx. 65°C were designed as described in the bioinformatics part of the methods section. PCR reactions were set up using the following 50 µl reaction mixture:

38.0 µl	Nuclease-free water
5.0 µl	10x *Pfx* amplification buffer (Invitrogen)
1.0 µl	50 mM $MgSO_4$
1.5 µl	dNTP mix (10 mM each)
0.5 µl	Platinum *Pfx* DNA polymerase (Invitrogen, 2.5 U/µl)
2.0 µl	Template DNA
2.0 µl	Primer mix (10 pmol/µl)

Thermocycling was performed using hot-start and touch-down conditions:

Step	**[°C]**	**[hh.mm.ss]**		**# cycles**
[Initial denaturation	95	00:02:00	]	1x
[Denaturation	94	00:00:15		
Annealing	60—>55	00:00:30		
Extension	68	1 min/kb	]	6x
[Denaturation	94	00:00:15		

Annealing	60	00:00:30		
Extension	68	1 min/kb	]	30x*
[Final extension	68	00:07:00	]	1x

*To minimize the number of cycles and correspondingly, the number of mutations introduced during PCR, samples of 16 µl were taken after extension in cycles 20 and 25 for determination of lowest cycle numbers necessary to obtain PCR products.

10 µl of all samples were analyzed by agarose gel electrophoresis, and the samples with the lowest cycle numbers showing a band were selected for adapter PCR as follows.

(b) Gateway adapter PCR for annealing of the full-length recombination sites

Gateway adapter primers (attB1_fwd and attB2_rev) with annealing temperatures of ~50°C were used according to the manufacturer's recommendations. PCR reactions were set up using the following 50 µl reaction mixture:

35.4 µl	Nuclease-free water
5.0 µl	10x *Pfx* amplification buffer (Invitrogen)
1.0 µl	50 mM $MgSO_4$
1.5 µl	dNTP mix (10 mM each)
0.5 µl	Platinum *Pfx* DNA polymerase (Invitrogen, 2.5 U/µl)
5.0 µl	Product of gene-specific PCR
1.6 µl	Adapter primer mix (25 pmol/µl each)

Thermocycling was performed using hot-start conditions:

Step	**[°C]**	**[hh.mm.ss]**		**# cycles**
[Initial denaturation	95	00:02:00	]	1x
[Denaturation	94	00:00:15		
Annealing	45	00:00:30		
Extension	68	1 min/kb	]	5x
[Denaturation	94	00:00:15		
Annealing	65	00:00:30		
Extension	68	1 min/kb	]	20x*
[Final extension	68	00:07:00	]	1x

* To minimize the number of cycles and correspondingly, the number of mutations introduced during PCR, samples of 16 µl were taken after extension in cycles 10 and 15 for determination of lowest cycle numbers necessary to obtain PCR products.

5 µl of all samples were analyzed by agarose gel electrophoresis, and the samples with the lowest cycle numbers showing a band were selected for PEG-8000 purification and recombinatorial cloning, as described below.

(c) Colony PCR for verification of cloned insert sizes

Vector-specific primers with annealing temperatures of ~55°C were used according to the constructs to be tested. PCR reactions were set up using the following 30 µl reaction mixture:

```
25.1 µl  Nuclease-free water
 3.0 µl  10x Taq amplification buffer*
 0.6 µl  dNTP mix (10 mM each)
 0.8 µl  Taq DNA polymerase (in-house, 10 U/µl)
 0.5 µl  Primer mix (25 pmol/µl each)

* 10x Taq amplification buffer: 650 mM Tris-Cl pH 8.8, 166 mM
(NH4)2SO4, 25 mM MgCl2, 0.1% (v/v) Tween-20, filtered through 0.45 µm
filter, aliquoted à 1 ml and stored at -20°C.
```

Single bacterial colonies were picked with 1 µl inoculation loops and transferred into the PCR mix by stirring the loop in the mix for about 2 sec. In parallel, 100 µl selective medium in 96-well plates were inoculated with the same inoculation loop for preparation of glycerol stocks. Plates were incubated at 37°C with shaking until analysis by colony PCR was finished.

Thermocycling was performed using hot-start conditions and a long initial denaturation to denature bacterial nucleases and other inhibitory proteins:

Step	**[°C]**	**[hh.mm.ss]**		**# cycles**
[Initial denaturation	94	00:05:00	]	1x
[Denaturation	94	00:00:30		
Annealing	50	00:00:30		
Extension	72	30 sec/kb	]	30x
[Final extension	68	00:07:00	]	1x

10 µl of all samples were analyzed by agarose gel electrophoresis. 96-well plate glycerol stocks of bacterial colonies showing the correct insert sizes were used for inoculation of 5 ml selective medium in culture tubes for minipreps of plasmid DNA.

2.1.7 Cloning of PCR products

PCR products with the lowest cycle numbers showing a band were selected for PEG-8000 purification to remove primer dimers of up to 100 bp length which result in high unspecific background during the following recombination reaction. The following protocol recommended for PCR purification by Invitrogen was used:

```
 10 µl   PCR product
+30 µl   1x TE
+20 µl   30% PEG-8000 with 30 mM MgCl2
```

The mixtures were vortexed immediately after the addition of PEG-8000/$MgCl_2$ and centrifuged with 16,000 g for 15 min at room temperature. Supernatants were removed quickly, and the DNA pellets were resuspended in 10 µl 1x TE.

Purified PCR products were used for recombinatorial cloning ('BP reaction') into Gateway entry vectors using components of Invitrogen's 'PCR Cloning System'. The following reaction set-up (scaled down from 20 µl to 10 µl in order to save expenses) was used:

```
4 µl   1x TE
2 µl   5x BP reaction buffer (Invitrogen)
1 µl   pDONR vector (Invitrogen, 150 ng/µl)
2 µl   Purified PCR product (approx. 50 fmol)
1 µl   BP clonase enzyme mix (Invitrogen)
```

The reaction mixtures were vortexted gently and collected at the bottom of the tubes by short centrifugation. The clonase reaction was performed in a thermocycler at 25°C with heated lid (30°C) for 4-16 h. BP clonase was inactivated by digestion with 2 µg Proteinase K at 37°C for 10 min. Proteinase K was then heat-inactivated at 95°C for 3 min. Samples were frozen and stored at -20°C until transformation.

BP reaction products were transformed into *E. coli* DH5α. Single colonies were used for inoculation of glycerol stocks in 96-well plates and for colony PCR. Plasmids showing the correct insert size were purified from 5 ml *E. coli* culture and used for DNA sequencing of the 5' and 3' ends of the insert to verify identity, reading frame and nucleotide sequence of the cloned ORFs.

2.1.8 DNA sequencing

In order to verify identity, reading frame and nucleotide sequence of the cloned ORFs, vector primers for pDONR201 (pDONR201_ins_fwd and pDONR201_ins_rev)

were used for DNA sequencing of short plasmid stretches and the adjacent 5' and 3' ends of the cloned ORFs. 200 ng of plasmid DNA were used for thermocycling with BigDye Terminator Mix (Applied Biosystems) according to the following protocol for 10 µl sequencing reactions in 8-strips for PCR:

```
0.5 µl   Plasmid DNA (ca. 200 ng)
```

was incubated at 96°C for 1 min to ensure proper denaturation and put onto ice. Then the sequencing mix containing the following components was added:

```
5.3 µl   Nuclease-free water
4.0 µl   BigDye Terminator Mix (Applied Biosystems)
0.2 µl   Primer (forward or reverse, 50 µM)
```

Thermocycling was performed using hot-start conditions using the following cycling parameters:

Step	**[°C]**	**[hh.mm.ss]**		**# cycles**
[Initial denaturation	96	00:03:00	]	1x
[Denaturation	96	00:00:20		
Annealing	45	00:00:10		
Extension	60	00:04:00	]	35x

Sequencing products were purified by ethanol precipitation according to the following protocol. 25 µl 96% ethanol (2.5 vol) were added to each 10 µl product and mixed by inverting the tubes five times. After centrifugation for 1 h with 2,700 g at room temperature, supernatants were discarded by inverting the strips onto paper towels. DNA pellets were washed two times by addition of 150 µl 70% ethanol and centrifugation for 30 min with 2,700 g at room temperature. The resulting pellets were dried in a vacuum centrifuge for 10 min with medium heat and stored at -20°C in the dark until determination of the sequence on an ABI PRISM Genetic Analyzer (Applied Biosystems).

2.1.9 Subcloning into expression vectors

Gateway entry clones with verified ORF inserts were selected for subcloning into various expression vectors. Recombinatorial cloning ('LR reaction') into Gateway destination vectors was performed with components supplied with Invitrogen's 'LR clonase enzyme mix'. The following reaction set-up (scaled down from 20 µl to 10 µl in order to save expenses) was used:

```
5.5 µl   1x TE
2.0 µl   5x LR reaction buffer (Invitrogen)
0.5 µl   pDEST vector (150 ng/µl)
1.0 µl   pENTR-ORF plasmid DNA (approx. 35 fmol)
1.0 µl   LR clonase enzyme mix (Invitrogen)
```

The reaction mixtures were vortexted gently and collected at the bottom of the tubes by short centrifugation. The clonase reaction was performed in a thermocycler at 25°C with heated lid (30°C) for 4-16 h. LR clonase was inactivated by digestion with 2 µg Proteinase K at 37°C for 10 min. Proteinase K was then heat-inactivated at 95°C for 3 min. Samples were frozen and stored at -20°C until transformation.

LR reaction products were transformed into *E. coli* DH5α. Single colonies were used for inoculation of glycerol stocks in 96-well plates and for colony PCR. Plasmids showing the correct insert size were purified from 5 ml *E. coli* culture.

2.1.10 Quantification of nucleic acids

Total nucleic acid quantity and quality of purified DNA and RNA was determined using a NanoDrop ND-1000 UV-Vis Scanning Spectrophotometer according to the manufacturer's recommendations. Nucleic acid concentrations corresponding to OD_{260}=1 were set to 50 µg/ml (dsDNA), 40 µg/ml (ssRNA), 33 µg/ml (ssDNA) and 20 µg/ml (oligonucleotides) at pH 7.

2.2 Methods in protein biochemistry

2.2.1 Denaturing polyacrylamide gel electrophoresis (SDS-PAGE)

For seperation of proteins using their mobility during denaturing gel electrophoresis, discontinuous SDS-PAGE according to Laemmli was used. Electrophoresis was performed in a Mini Protean III electrophoresis cell (BioRad) according to the manufacturer's recommendations. Stock solutions used for gel preparation were:

```
1.5 M Tris-Cl, pH 8.8 (for seperating gel)
0.5 M Tris-Cl, pH 6.8 (for stacking gel)
10 % (w/v) sodium dodecyl sulfate (SDS)
30 % acrylamide/0.8 % N'N'-bis-methylene acrylamide
       (BioRad)
10 % ammonium peroxodisulfate (APS)
N,N,N',N'-Tetramethylethylenediamine (TEMED)
```

Seperating gels were prepared immediately before use and contained 375 mM Tris-Cl pH 8.8, 0.1 % (w/v) SDS, 7.5–20 % (w/v) acrylamide/bisacrylamide, 0.05 % (w/v) APS and 0.05 % (v/v) TEMED. Stacking gels poured on top of the seperating gels contained 125 mM Tris-Cl pH 6.8, 0.1 % (w/v) SDS, 5 % (w/v) acrylamide/bisacrylamide, 0.1 % (w/v) APS and 0.2 % (v/v) TEMED. Protein samples were diluted in 3x SDS-PAGE sample buffer containing 94 mM Tris-Cl pH 6.8, 30% (v/v) glycerol, 3% (w/v) SDS, 0.02% (w/v) bromophenol blue and 100 mM dithiothreitol as reducing agent. Electrophoresis was performed in SDS-PAGE running buffer composed of 0.25 M Tris-Cl, 1.92 M glycine and 1% (w/v) SDS, pH 8.3.

2.2.2 Coomassie staining of polyacrylamide gels

For staining of proteins seperated by SDS-PAGE, polyacrylamide gels were incubated for ≥1 h in Coomassie staining solution containing 40% (v/v) methanol, 7% (v/v) acetic acid and 0.1% (w/v) Coomassie Brilliant Blue R250. Then, gels were destained 2-3 times for 30 min in destaining solution containing 20% (v/v) methanol and 7% (v/v) acetic acid until protein bands were clearly distinguishable from background staining.

2.2.3 Western blotting

Transfer of proteins from SDS-PAGE gels to nitrocellulose membranes was performed in a Mini Trans-blot electrophoretic transfer cell (BioRad) according to the manufacturer's recommendations. The transfer buffer contained 25 mM Tris-Cl pH 8.3, 192 mM glycine and 20% (v/v) methanol. Upon completion of Western blotting, nitrocellulose membranes were incubated for 5 min in Ponceau S solution (0.1% (w/v) Ponceau S in 5% acetic acid) to visualize transferred proteins. Afterwards, membranes were destained with PBS and either immediately blocked (see below), stored in PBS at 4°C for up to three days, or dried and stored at -20°C until further use.

2.2.4 Immunodetection of proteins

For visualization of endogenous as well as recombinant proteins after seperation by SDS-PAGE and Western blot transfer onto nitrocellulose membranes, immunodetection was performed using target-specific primary antibodies in combination with antibody-specific secondary antibodies for enzymatic detection.

Choice of blocking proteins and buffer components

Blocking of unspecific protein binding sites in the nitrocellulose membrane was carried out with BSA or non-fat milk depending on which gave better results in the specific experimental set-up. In most cases, phosphate buffered saline (PBS) was used for incubation steps, whereas Tris buffered saline (TBS) was used when the secondary antibody was coupled to alkaline phosphatase, since the phosphate in PBS can result in product inhibition of the phosphatase reaction used for enzymatic detection.

Two different protocols were established for monoclonal mouse IgG and polyclonal rabbit IgG antibodies:

(a) Immunodetection using monoclonal mouse IgG

Blocking: 1 h to overnight at 4°C in PBS pH 7.4 with 3% (w/v) BSA (Fraction V, ≥99% purity)

Primary antibody: 2 h at room temperature with monoclonal mouse anti-HA IgG (Sigma), diluted 1:2,500 in PBS pH 7.4 with 1% (w/v) BSA and 0.05% Tween-20

Washing: 3x 5 min in PBS pH 7.4 with 0.05% Tween-20

Secondary antibody: 1 h at room temperature with polyclonal sheep anti-mouse IgG, HRP conjugate (Amersham), diluted 1:10,000 in PBS pH 7.4 with 1% (w/v) BSA and 0.05% Tween-20

Washing: 3x 5 min in PBS pH 7.4 with 0.05% Tween-20

Detection: 2 ml / blot of ECL substrate (Amersham), exposure of Hybond ECL film (Amersham) for different times to obtain at least one film with good signal to background ratio.

(b) Immunodetection using polyclonal rabbit IgG

Blocking: 1 h to overnight at 4°C in PBS pH 7.4 with 3% (w/v) BSA

Primary antibody: 2 h at room temperature with polyclonal rabbit anti-FLAG IgG (Sigma), diluted 1:2,500 in PBS pH 7.4 with 1% (w/v) BSA and 0.05% Tween-20

Washing: 3x 5 min in PBS pH 7.4 with 0.05% Tween-20

Secondary antibody: 1 h at room temperature with polyclonal donkey anti-rabbit IgG, HRP conjugate (Amersham), diluted 1:10,000 in PBS pH 7.4 with 1% (w/v) BSA and 0.05% Tween-20

Washing: 3x 5 min in PBS pH 7.4 with 0.05% Tween-20

Detection: 2 ml / blot of ECL substrate (Amersham), exposure of Hybond ECL film (Amersham) for different times to obtain at least one film with good signal to background ratio.

2.2.5 Expression of recombinant proteins in *E. coli*

E. coli expression strain Rosetta(DE3)pLysS (Novagen) containing a T7 RNA polymerase gene under control of the IPTG-inducible *lac* promoter was transformed with pEXP17-ORF expression plasmids harboring cloned Hsa21 ORFs downstream of a T7 promoter and an encoded N-terminal hexahistidine tag. After transformation, bacterial colonies were grown overnight at 37°C on LB-agar plates containing 100 µg/ml Ampicillin (for selection of transformants), 12.5 µg/ml Chloramphenicol (for maintenance of the pRARE plasmid) and 1% glucose (for suppression of background protein expression).

Single bacterial colonies were picked for insert amplification by colony PCR. Colonies with correct insert sizes (usually >95% of all colonies) were used for preparation of glycerol stocks in 96-well plates and for inoculation of liquid pre-cultures (2-5 ml LB with 100 µg/ml Ampicillin, 12.5 µg/ml Chloramphenicol and 1% glucose). Cells were grown at 37°C with 250 rpm shaking until optical density at 595 nm reached about 0.6.

Expression for protein solubility test (high-throughput protocol)

For larger numbers of solubility tests, protein expression was performed in 96-well cell culture plates. 150 µl LB medium with 100 µg/ml Ampicillin, 12.5 µg/ml Chloramphenicol and 0.1% (w/v) glucose were inoculated from 96-well glycerol stocks using pin replicators. After growth at 37°C with 250 rpm until cultures reached OD_{595}≈1.0 (usually 6-8 h), protein expression was induced by addition of IPTG to a final concentration of 0.1 mM. Expression was performed for 10 h at 30°C with 250 rpm. Expression cultures were immediately used for the high-throughput protein solubility test described below.

2.2.6 Determination of protein solubility

For solubility testing of 96 expression cultures in parallel, a commercial kit was used (Relay 96 Protein Screen, Invitrogen). 150 µl expression cultures were transferred into 96-well filter plates, and 50 µl native lysis buffer (supplied with the kit) containing 25 µg/ml lysozyme was added. Cell lysis was performed for 15 min at room temperature. The soluble fraction was filtered through the plate bottom into a receiver plate by centrifugation for 5 min with 2,500 g. The remaining insoluble fractions were solubilized by addition of 200 µl denaturing buffer (supplied with the kit) and incubation for 10 min at room temperature. This fraction was then also filtered through the plate bottom into a receiver plate by centrifugation for 5 min with 2,500 g. 30 µl of each soluble and insoluble fraction was added to 15 µl 3x SDS-PAGE sample buffer, incubated for 10 min at 95°C and frozen at -20°C until further analysis.

2.2.7 Quantification of proteins

Protein concentrations were quantified using the Bradford assay, in which a differential color change of a dye occurs in response to various concentrations of protein. The absorbance maximum for an acidic solution of Coomassie Brilliant Blue

G-250 dye shifts from 465 nm to 595 nm upon binding primarily to basic and aromatic amino acid residues, especially arginine.

Ovalbumin was used for standard curves, since the use of bovine serum albumin results in underestimation of the concentrations of other proteins due to its relatively strong binding of the Coomassie dye. Concentrations of detergent higher than 0.1% and high pH levels can interfere significantly with this assay. When detergent was present in the samples (e.g. in COS-1 cell lysate for coimmunoprecipitation, see below), it was partially compensated for by preparing the ovalbumin solution for the standard curve with the very same buffer.

Concentrations were measured with BioRad Protein Assay using standards of 20, 10, 5, 2.5, 1.25 and 0 µg/ml ovalbumin in ddH_2O. Protein samples were diluted 1:10 and 1:100 in a final volume of 800 µl ddH_2O. After addition of 200 µl BioRad dye reagent, samples were mixed well and incubated for 10 min at room temperature. The absorbance at 595 nm was measured in 1 cm plastic cuvettes, and protein concentrations were determined by comparison with the concentration of ovalbumin at the same value of absorption.

2.3 Methods in cell biology

2.3.1 Transformation of yeast cells

Fresh yeast cells L40ccua were grown in YPD medium to OD_{600} ≈ 0.6-0.8. Cells were washed once in 0.33 volumes of 10 mM Tris pH 7.4 with 1 mM EDTA and resuspended in 0.03 volumes of 5 mM Tris pH 7.4 with 100 mM lithium acetate, 1 M sorbitol and 0.5 mM EDTA. After 10 min incubation at room temperature, 10 µl of cell suspension was mixed with 125 ng of each plasmid, 12.5 µg salmon sperm DNA and 55 µl of 10 mM Tris pH 7.4 with 50 mM lithium acetate, 40% PEG-3350 and 1 mM EDTA. After incubation for 30 min at 30°C, 7.5 µl DMSO were added. Heat shock was performed for 7 min at 42°C, cells were pelleted with 2,000 g for 5 min and resuspended in 100 µl water. After plating onto SD2 plates, colonies were grown for 3 days at 30°C.

2.3.2 Autoactivation test

Bait constructs were transformed into yeast strain L40ccua and mated with yeast strain L40ccα containing the empty prey vector pGAD426-D3. Diploid cells were transferred onto minimal medium plates lacking tryptophane, leucine, histidine, and uracil (SD4 medium) with and without nylon membranes and incubated for five days at 30°C. Autoactivation of baits was determined by growth on SD4 minimal medium and by a β-galactosidase (*LacZ)* assay.

For the β-galactosidase assay, nylon membranes with yeast colonies were frozen for 1 min in liquid nitrogen to disrupt the yeast cells. Then, the membranes were incubated for 4 h at 37°C on Whatman paper soaked in X-Gal mix. For this mix, 10 ml Z buffer (10.7 g/l Na_2HPO_4*2 H_2O, 5.5 g/l NaH_2PO_4*1 H_2O, 0.75 g/l KCl, 0.246 g/l $MgSO_4$*7 H_2O, pH 7) were supplemented with 156 µl of 2% (w/v) X-Gal [5-bromo-4-chloro-3-indolyl-D-galactoside in N,N-dimethylformamide] and 100 µl of 1 M DTT. Stained membranes were dried on Whatman paper overnight in the dark.

2.3.3 Yeast two-hybrid (Y2H) assays

Two different yeast two-hybrid set-ups were used during this work. For the identification of new interactions involving Hsa21 proteins, an automated high-throughput Y2H mating array screen was performed in collaboration with the group of

Prof. Erich Wanker at the Max Delbrück Center for Molecular Medicine (Berlin). For verification of the identified interactions and also for testing of potential interologs, a low-throughput two-hybrid assay after yeast cotransformation was conducted.

Mating-array Y2H screen

Bait constructs containing Hsa21 ORFs in fusion with the LexA DNA-binding domain enabled growth on trp^- medium. A Y2H array containing 5,640 prey clones was screened. 3,589 prey clones were derived from the human fetal brain expression library hEx1 (Bussow *et al.* 1998), and 2,051 prey clones were product of a Gateway recombinational cloning approach that was used to shuttle full-length human ORFs from publicly available Gateway entry vectors (German Resource Center for Genome Research, RZPD) into Y2H plasmids. Prey constructs contained the prey ORFs in fusion with the GAL4 activation domain and enabled growth on leu^- medium.

Plasmids encoding baits and preys were transformed into yeast strains L40ccua and L40ccα, respectively. L40ccα clones were arrayed into 96-well microtiter plates and mixed with pools of six L40ccua clones for interaction mating. Diploid cells were transferred onto YPD medium plates and, after incubation for 24 h at 30°C, onto plates with minimal medium lacking tryptophane and leucine (SD2 medium) for an additional 72 h at 30°C. For two-hybrid selection, diploid cells were transferred onto SD4 minimal medium plates with and without nylon membranes and incubated for 5 days at 30°C. The nylon membranes were subjected to the β-*galactosidase* (*LacZ*) assay. Positive pools were screened again as single L40ccua clones to identify the interacting bait.

Two-hybrid assay after cotransformation

Fresh L40ccua yeast cells were grown in YPD medium to an optical density at 600 nm of 0.6-0.8. Cells were washed once in 0.33 volumes of 10 mM Tris-Cl pH 7.4 with 1 mM EDTA and resuspended in 0.03 volumes of 5 mM Tris pH 7.4 with 100 mM lithium acetate, 1 M sorbitol and 0.5 mM EDTA. After 10 min incubation at room temperature, 10 µl of this cell suspension was mixed with 125 ng of each plasmid, 12.5 µg salmon sperm DNA and 55 µl 10 mM Tris pH 7.4 with 50 mM lithium acetate, 40% PEG-3350 and 1 mM EDTA. After incubation for 30 min at 30°C, 7.5 µl DMSO were added. Heat shock was performed for 7 min at 42°C. Cells were collected by centrifugation at 2,000 g for 5 min and resuspended in 100 µl distilled water. After

plating onto SD2 plates, colonies were grown for 3 days at 30°C. Three independent L40ccua colonies from each co-transformation were arrayed into 96-well plates and replicated onto one SD2 plate for growth monitoring, one SD2 plate which was then covered with a nylon membrane for LacZ assay, and one SD4 plate for the auxotrophy reporter assay. All plates were incubated for three days until colonies were visible. The nylon membranes were subjected to a β-galactosidase (LacZ) assay.

2.3.4 Culture of COS-1 cells

The culture conditions used here followed the recommendations of the American Type Culture Collection (ATCC).

Propagation

The adherent mammalian kidney fibroblast-like cell line COS-1 (derived from the african green monkey *Cercopithecus aethiops*) was propagated in Dulbecco's modified Eagle's medium (Gibco, DMEM with 1 g/l glucose, L-glutamine and pyruvate) supplemented with 10% (v/v) fetal bovine serum (FBS, Gibco, heat-inactivated) and 1:100 Penicillin/Streptomycin (Gibco, 10 kU/ml Penicillin G sodium and 10 mg/ml Streptomycin sulfate) at a temperature of 37°C in an atmosphere of 95% air and 5% CO_2.

Subculturing

When the cells in a 75 cm^2 culture flask reached about 80% confluence level (usually 2-3 times/week), the medium was removed and discarded, and the cell layer was rinsed briefly with 10 ml PBS (PanBiotech, cell culture grade, without magnesium and calcium) to remove all traces of serum which contains trypsin inhibitor. 1 ml of Trypsin-EDTA solution (GIBCO, 0.25% Trypsin with 1 mM EDTA•4Na) was added, and cells were detached at 37°C for 2-4 min, as monitored by microscopy. Cells were aspirated in 6 ml of complete medium and added to new culture vessels in a sub-cultivation ratio of 1:6.

Preservation of cells

For storage, 2-3 x 10^6 cells were frozen in complete growth medium (without antibiotics) containing 5% DMSO (Merck, cell culture grade). Storage temperature

was liquid nitrogen vapor temperature. After thawing in warm water, the cell suspension was quickly diluted in warm complete growth medium and added to new culture vessels in a subcultivation ratio of 1:2.

2.3.5 Transient transfection of COS-1

The transfection protocol applied here uses lipopolyamine complexes for transient transfection of mammalian cells with expression plasmids. The following protocol describes transfection in 12-well plates (4 cm^2 surface area / well). If other plate or flask formats were used, the amounts of cells, DNA and solutions was scaled up or down according to the surface area.

Transfection with DreamFect reagent

In 12-well plates, 1 x 10^5 cell were plated onto cover slips in 1 ml complete medium with antibiotics per well, so that the cells were about 80% confluent 24 h later at the time of transfection. 1.6 µg plasmid DNA (for double transfections, 800 ng of each plasmid) were diluted in 50 µl DMEM, and 4 µl DreamFect (Ozbiosciences) were diluted in 50 µl DMEM. Both solutions were combined within 5 min, and incubated for 20 min at room temperature. 100 µl complexes were added to each well, and after gentle mixing, the plates were incubated for 24-36 h at 37°C with 5% CO_2 prior to testing for protein expression.

2.3.6 Immunolocalization of proteins

After transient transfection and incubation for 24-36 hours, cells were washed twice with cold PBS and fixed for 10 min with 3.7% paraformaldehyde in PBS. Cells were permeabilized for 10 min with PBS containing 0.1% Triton X-100, washed twice with PBS and blocked for 60 min in PBS supplemented with 3% BSA. After washing with PBS containing 0.1% Triton X-100, the cover slips were incubated for 60 min at room temperature with specific antibodies directed against FLAG (rabbit polyclonal anti-FLAG IgG from Sigma, 1:100 in PBS with 0.1% Triton X-100) and HA (monoclonal anti-HA IgG from Sigma, clone HA-7, 1:100 in PBS with 0.1% Triton X-100). After 3x5 min washing in PBS with 0.1% Triton X-100, the cover slips were incubated for 60 min at room temperature with PBS containing the detection antibodies (1:300 anti-Rabbit-Cy3 conjugate and 1:150 anti-mouse-FITC conjugate, both from Dianova). After 2x5 min washing with PBS containing 0.1% Triton X-100, nuclei were stained

with DAPI (Sigma), and an anti-fade mounting medium (VectaShield, Vector Labs.) was used to prepare the cover slips for microscopy. Analysis of subcellular localizations was performed using a confocal fluorescence microscope (LSM 510, Zeiss), and image analysis was carried out with the LSM 510 software version 4.0 SP2.

2.3.7 Pull-down assay

Expression of recombinant proteins in COS-1

For expression of FLAG- and hemagglutinin (HA)-tagged fusion proteins, COS-1 cells were seeded in 75 cm^2 flasks and grown for 24 h to reach cell confluence of 70–80%. Transient transfection was performed with 20 µg of plasmid DNA encoding FLAG- or HA-tagged proteins using DreamFect reagent as described above. After 24h, cells were washed twice with cold PBS and incubated, with occasional shaking, with 800 µl cold lysis buffer containing 50 mM HEPES pH 7.4 with 150 mM NaCl, 1.5 mM $MgCl_2$, 1 mM DTT, 5 mM imidazole, 10% glycerol, 1% NP-40, 0.2 mM EDTA, 20 mM NaF, 1x Complete EDTA-free protease inhibitor cocktail (Roche) and 30 U/ml Benzonase (Sigma). Cellular debris was removed by centrifugation for 10 min with 10,000 g at 4°C. Protein concentrations of the cleared lysates were determined by Bradford assay (BioRad) and adjusted to 1 µg/µl with lysis buffer. The normalized lysates were used for pull-down, as described below.

Binding of recombinant proteins from *E. coli* to Ni-NTA beads

For production of hexahistidine (6xHis) fusion proteins, ORFs were subcloned into pDEST17 vector (Invitrogen) and expressed in *E. coli* Rosetta(DE3)pLysS (Novagen). Cells were washed with PBS and resuspended in lysis buffer containing 50 mM Tris-Cl pH 8.0, 4 mM $MgCl_2$, 5 mM imidazole, 1% Triton X-100, 0.5% NP-40, 1 mg/ml lysozyme, 1x Complete EDTA-free protease inhibitor cocktail (Roche) and 30 U/ml Benzonase (Sigma). After incubation for 30 min on ice and sonication for 1 min at 4°C in a water bath sonicator, cellular debris was removed by centrifugation with 10,000 g for 30 min at 4°C. The supernatants containing 6xHis fusion proteins were added to 50 µl Ni-NTA agarose beads (Qiagen) preblocked with BSA. Control beads were incubated with supernatant without 6xHis protein. After binding for 60 min on a wheel at 4°C, beads were washed with 3x 300 µl cold wash buffer (50 mM

Tris-Cl pH 7.4 with 400 mM NaCl and 5 mM imidazole) and used for pull-down of FLAG- or HA-tagged proteins from COS-1 lysate.

Pull-down from COS-1 lysates using 6xHis fusion proteins on Ni-NTA beads

COS-1 lysates containing 400 µg total protein were incubated with control beads and with 6xHis fusion protein (~50 µg) bound to 50 µl beads for 60 min at 4°C on a wheel. Beads were washed five times with 300 µl cold binding buffer containing 50 mM HEPES pH 7.4 with 300 mM NaCl, 25 mM imidazole, 1 mM DTT, 10% glycerol, 1% NP-40, 0.5% Triton X-100, 20 mM NaF and 1x Complete EDTA-free protease inhibitor cocktail (Roche). Then, the beads were collected, transferred to new 1.5 ml tubes and boiled in 3x Laemmli sample buffer for 10 min at 95°C with shaking. Eluted proteins were analyzed by SDS-PAGE and immunoblotting with anti-HA antibody (Sigma, clone HA-7) or polyclonal anti-FLAG antibody (Sigma). Proteins were visualized by chemoluminescence detection using SuperSignal West Femto reagent (Pierce).

2.4 Computational methods

2.4.1 ORF primer design

PCR primer sequences were selected using PRIDE, a primer design program that automatically designs primers for large scale cloning and sequencing projects (Haas *et al.* 1998). The ORF-specific part of the 5' (forward) amplification primer was fixed to include the ATG start codon. The 3' (reverse) amplification primer was designed to an appropriate sequence in the 3' untranslated region immediately downstream of the stop codon of the transcript used as reference sequence. Annealing temperature was considered optimal in the range of 63 to 67°C. Then, short Gateway attB overhangs of 9 and 11 bp for forward and reverse primers, respectively, were added to the sequences.

2.4.2 Conversion of accession numbers to Entrez GeneIDs

The following strategy, combining the conversion power of three independent identifier conversion tools, permitted the ID conversion for nearly all proteins analyzed. GI and protein accession numbers describing proteins participating in PPIs were used as input for Ensembl's BioMart (release 34) data mining tool (Hubbard *et al.* 2007) to retrieve the corresponding HGNC gene symbols and Entrez GeneIDs. Those accession numbers not resolved by BioMart were used as input for the Nome Protein Identification Resolution Database (http://pathogene.swmed.edu/nome/). Accession numbers still not resolved were used as input for the DAVID Gene ID Conversion Tool (Dennis *et al.* 2003).

2.4.3 Retrieval of orthologous interaction data (interolog data)

Orthology relationships between human proteins and proteins from *S. cerevisiae*, *D. melanogaster* and *C. elegans* were retrieved from the InParanoid database (version 4) of pair-wise orthologs (Remm *et al.* 2001). Here, orthology predictions were available for 2,171 clusters in yeast, 5,610 in fly and 4,599 in worm. From these predictions, all chromosome 21-encoded human proteins were extracted using Ensembl GeneIDs for all Hsa21 genes. Thus, 49 Hsa21 protein orthologs were found in yeast, 105 orthologs in fly and 94 orthologs in worm (Supplemental Table S1, see http://chr21.molgen.mpg.de/21ppi/S1_orthologs.xls). The Hsa21 orthologs were used

to retrieve all corresponding interaction data from large-scale Y2H screens previously performed for the three model organisms (Schwikowski *et al.* 2000; Giot *et al.* 2003; Li *et al.* 2004b).

As result, 10 interactions for six Hsa21 orthologs were found in yeast, 62 interactions for 25 orthologs in fly and 57 interactions for seven orthologs in worm. However, 45 of the 57 worm interologs involved a 'sticky' protein, C03A7.4, ortholog of the keratin associated protein KRTAP10-10, and were not analyzed further. In total, the resulting Y2H sets of orthologous interactions (so-called "interologs") consisted of 84 interactions involving 31 Hsa21 orthologs.

In addition, using systematic literature searches in PubMed, another 51 interologs stemming from nine different species were identified. Taken together, the interolog collection for Hsa21 consisted of 135 interologs for 50 different Hsa21 orthologs (Supplemental Table S2, see http://chr21.molgen.mpg.de/21ppi/S2_interologs.xls).

2.4.4 Retrieval of known protein-protein interactions from literature-curated databases

The complete BIND data set was obtained from the BIND FTP site (20050811.nrints.txt). The primary file contained a non-redundant list of 77,929 interactions in the BIND database. Only the interactions of type 'protein' and taxon '9606' (human) were extracted to receive a list of 5,420 human protein-protein interactions, excluding interaction data for small molecules, DNA, RNA and complexes. Next, 601 interactions were removed between proteins and small peptides characterized by PDB structure accession numbers. For the remaining 4,819 interactions, the GI numbers from the primary file were converted to HGNC gene symbols and Entrez GeneIDs, as described above. HGNC gene symbols and GeneIDs could be obtained for 4,699 human PPIs between 2,792 different proteins from BIND.

The complete HPRD data set was obtained from the HPRD FTP site (psimi_single_final_09_13_05.tar.gz). This XML file contained a list of 20,112 interactions in the HPRD database. All interactions with two partners were extracted, resulting in a list of 19,514 protein-protein interactions, excluding 598 complexes. Then, the corresponding HGNC gene symbols, Entrez GeneIDs and TaxonIDs were extracted from the XML file. Next, 10,367 interactions with non-human partners and

1,055 PPIs without GeneID were removed to obtain a set of 7,705 human PPIs between 4,440 different proteins from the HPRD database.

2.4.5 Retrieval of known protein-protein interactions from previous large-scale Y2H screens

The data set of protein interactions from the MDC Y2H screen was obtained from Supplemental Table S4 accompanying the original publication (Stelzl *et al.* 2005). This file contained a list of 3,186 interactions. The HGNC symbols and Entrez GeneIDs were extracted for all proteins, then 27 interactions with non-human partners and 9 PPIs with ambiguous identifiers were removed to obtain a set of 3,149 human PPIs between 1,682 different proteins.

The data set of protein interactions from the CCSB Y2H screen was obtained from Supplemental Table S2 accompanying the original publication (Rual *et al.* 2005). This file contained a list of 2,754 interactions. The HGNC symbols and Entrez GeneID were extracted for all proteins to obtain a set of 2,754 human PPIs between 1,549 different proteins.

2.4.6 Retrieval of known protein-protein interactions from PubMed

The PubMed database (http://www.ncbi.nlm.nih.gov) was manually searched for known interactions of all 284 proteins encoded on Hsa21. Full gene descriptions as well as all synonymous gene symbols in Entrez Gene (Maglott *et al.* 2005) were used to identify publications describing known protein interactions. For each gene, a strategy of four consecutive steps was applied to obtain a maximum number of interactions for Hsa21:

(1) Collection of all synonymous gene/protein symbols and names from our own catalog and from NCBI's Entrez Gene database

(2) Retrieval of all publications containing the gene/protein names in title or abstract

(3) Restriction to publications reporting interaction studies, using the boolean expression "AND(interacting OR interact OR interacts OR interaction OR binding OR bind OR binds OR association OR associates OR associate)"

(4) manual curation to identify actual physical protein-protein interactions. This search resulted in a list of 303 interactions for 71 different Hsa21 proteins.

2.4.7 Visualization of protein interaction networks

Protein-protein interaction networks were visualized using the open source software Cytoscape 2.5.1, which is freely available at http://www.cytoscape.org/. Cytoscape is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating these interactions with gene expression profiles and functional data.

SIF format

For input of molecular interaction data, raw interaction files (SIF format) were created containing lists of protein-protein interaction pairs. Lines in the SIF file specify a source node, a relationship type (or edge type), and one or more target nodes:

```
nodeA <relationship type> nodeB
nodeC <relationship type> nodeA
nodeD <relationship type> nodeE
...
nodeY <relationship type> nodeZ
```

Every node and edge in Cytoscape has an identifying name, most commonly used with the node and edge data attribute structures. All node names were checked to be unique, as only identically named nodes will be treated as identical nodes.

The tag <relationship type> can be any string, but for the interaction data analyzed here, the common interaction type “pp” used in the Systems Biology community for protein-protein interactions was used. Tab characters were used to delimit the names in the simple interaction file format. Networks in simple interactions format were stored in files with a .sif extension, which are recognized by Cytoscape as SIF files.

Node and edge attributes

Interaction networks are most powerful when integrated with additional information. Cytoscape allows the user to add arbitrary node, edge and network information as node/edge/network attributes. These attributes can then be visualized in a user-defined way by setting up a mapping from data attributes to visual attributes (colors, shapes, etc.).

For input of attributes on nodes, node attribute files were formatted in the following way and stored using the suffix ".noa":

```
MolecularFunction
AHR = Trancription
```

```
ARNT = Trancription
ATF1 = Trancription
ATF2 = Trancription
...
ZMYND11 = other
```

2.4.8 Analysis of protein network topology

Topology analysis of protein-protein interaction networks was carried out using the 'NetworkAnalyzer' plug-in (version 1.0) in Cytoscape 2.5.1. This plugin computes specific parameters describing the network topology, including number of nodes, number of edges, number of self-loops, average number of neighbors, network diameter and average path length. More information can be found online at http://cytoscape.org/plugins2.php.

2.4.9 Retrieval of Gene Ontology (GO) annotations

GO annotations were obtained using the Babelomics FatiGO tool (Al-Shahrour *et al.* 2005) and Ensembl's BioMart (release 34) data mining tool (Hubbard *et al.* 2007) with Entrez GeneIDs as input. The combined conversion capabilities of two independent tools permitted retrieval of GO annotations for a maximum number of proteins. To extract proteins with GO annotations describing transcriptional regulator activity, all genes associated with the GO term GO:0030528 ('transcription regulator activity') were used.

2.4.10 Retrieval of pathway connections through PPI data

Retrieval of Hsa21 pathway annotation data

Pathway connections were obtained using the ArrayAnalyzer tool of the TransPath database, release professional 6.3 containing 11,330 genes and 57 pathways (Krull *et al.* 2006) with Entrez GeneIDs as input. 272 Hsa21 Entrez GeneIDs and 547 Entrez GeneIDs of direct interactors identified 389 molecules, which were used for subsequent analysis of functional groups. The cut-off for the probability of getting the observed number of hits in a pathway by random chance was set to $p \leq 0.05$ (calculated by the ArrayAnalyzer tool using a binomial distribution). The minimum number of hits from the data set in the pathway was set to n=2. As a result, nine distinct signal transduction pathways were found significantly associated with five

Hsa21 gene products and 59 direct interactors. Using the Hsa21 PPI data set, connections were established between a total of 35 Hsa21 proteins via 136 PPIs to 59 pathway-annotated interactors. All pathway connection data can be found in Supplemental Table S4 (see http://chr21.molgen.mpg.de/21ppi/S4_pathways.xls).

Sorting of pathways and proteins according to connection patterns

To differentiate pathway connections of membrane proteins and nuclear transcription factors from other proteins, the Hsa21 proteins were sorted according to their pattern of connections to different pathways using the wCLUTO data clustering application available online at http://cluto.ccgb.umn.edu/ (Rasmussen *et al.* 2003) with the following parameters: Cluster method – agglomerative; Number of clusters – 6, Similarity function – Correlation coefficient; Merging function – Group average; Value scaling – None; Cluster columns – true. This resulted in partial seperation of membrane proteins, cytosolic proteins and nuclear transcription factors, as visible in the pathway connection matrix.