2 Materials and Methods

2.1 PDZ Domain

2.1.1 PDZ Domain and Peptide Motif Nomenclature

Homologous residue positions are numbered according to a general PDZ domain numbering scheme (based on Songyang *et al.* [30]) and modified to account for the PDZ domain secondary structure elements (SSE) consensus. The residue number (e.g. αB:8) is composed of a secondary structure prefix (αB = second conserved alpha helix) and a number indicating the position of this residue within the conserved SSE (8 = 8th residue). Highly conserved sequence motifs outside SSEs are named separately, e.g. GLGF:2 indicates the L residue within the GLGF-loop.

All peptide sequence motifs are given according to a modified Seefeld Convention 2001 nomenclature [105]. 'x' denoting 'any amino acid', 'b' a mixture of 17 L-amino acids (without C/M/W), Φ representing hydrophobic amino acids, ψ aliphatic amino acids and π small amino acids. Deviating from this convention, we omitted the *fn/fc*-nomenclature and added $_{COOH}$ to a sequence to describe the free C-terminal carboxyl group and $_{CONH}$ for the C-terminal carboxamide group. All amino acids are given in the one letter code. Within a target sequence pattern, the different amino acids of one position are ordered due to their frequency of occurrence.

2.1.2 PDZ Domain Plasmid

DNA fragments encoding residues 985 to 1086 of AF6 and 1280 to 1371 of ERBIN were cloned into a pGEX6p2 expression vector (Amersham Biosiences, Freiburg, Germany). This expression vector allows growth selection due to its ampicillin resistance. It produces a N-terminal glutathione-S-transferase (GST)-tagged protein with a PreScission™ cleavage side. Both clones are kindly provided by Prof. Dr. K. Moelling (Institute of Medical Virology, Zurich, Switzerland). DNA fragments encoding residues 81 to 164 of SNA1 were introduced into a pGAT2 expression vector [106] with ampecillin resistance.

The expressed protein is N-terminal GST-tagged and C-terminal histidine (His)-tagged, enabling purification with two different columns.

For the transformation, *Escherichia coli* BL21 (DE3) competent cells (Stratagene Europe, Amsterdam, The Netherlands) were prepared using the calcium phosphate method [107]. The cells were thawed on ice for 30 min, after which 1 µg plasmid / 200 µl cell suspensions was added and left on ice for a further 30 min. The cell suspensions were incubated for exactly 90 s at 42°C and put back on ice for max. 5 min. After the addition of 400 - 600 µl LB media (without antibiotics) / 200 µl cell suspensions, the whole solution was incubated for 1 - 2 h at 37°C. For the selection of the cells incorporating the plasmid, the cell suspensions were struck out on an LB plate containing ampicillin and incubated over night at 37°C.

To obtain the best clone, with the highest expression level, it is useful to test the expression of some chosen clones. From the over night LB plate of the plasmid transformation, some clones were picked and grown in 2 ml LB media for 3 h at 37°C. After the addition of isopropyl β-D-thiogalacto-pyranoside (IPTG), the bacteria were left to grow for another three hours. One of the 2 ml cultures was left without IPTG and served as negative control. To determine which clone shows the best expression pattern, 1 ml from each culture was analyzed by SDS-PAGE [107]. The expression pattern could be optimized by the variation of IPTG concentration, of temperature and/or of expression time. To conserve the clone showing the optimal expression pattern, a 30 % glycerol stock-culture was made and stored at -80°C. (All compounds of the used media were described in [107].)

2.1.3 PDZ Domain Expression

E. coli BL21 (DE3) cells carrying the appropriate expression plasmid were grown at 37°C in LB medium containing 50 mg/l ampecillin for all three expression vectors to mid-log phase (OD_{600} = 0.5). Protein expression was induced by addition of IPTG under the optimal growth condition (Table 2.1). The cells were harvested by centrifugation at 7800 g for 10 min at 4°C. The pellet was washed and resuspended in 20 mM Tris buffer, Complete® (protease inhibitor cocktail tablets, Roche, Mannheim, Germany) pH 8.0 and stored at -80°C.

For the nuclear magnetic resonance (NMR) measurement, the AF6-PDZ protein was expressed in M9 medium with $^{15}NH_4Cl$ for a ^{15}N-labeled sample or $^{15}NH_4Cl/^{13}C$-glucose supplemented for a $^{15}N/^{13}C$-labeled sample under the same conditions. (All compounds of the used media were described in [107].)

Table 2.1 Optimized Expression Conditions.

Protein	IPTG [µM]	E_{600}	T [°C]#	t [h]	Cell [g/l]*	FS [mg/l]*	Protein [mg/l]*
AF6	0.5	0.5	28	3 - 4	~ 6	50 - 70	15-20
ERBIN	0.5	0.4	30	4 - 5	~ 9	60 - 80	n.d.
SYNA1	1.0	0.6	37	2 - 3	~ 4	18 - 20	n.d.

Footnotes: FS = fusionprotein; # = temperature after IPTG induction; * = all weights are given per one liter cell culture

All three PDZ domains are successfully expressed after small ameliorations of the growth conditions (Table 2.1). The AF6 and ERBIN PDZ domain expression using the pGEX-6p2 vector yielded approximately 50 - 80 mg/l fusion protein. 15 - 20 mg/ml pure AF6 PDZ domain was obtained after cleavage of the GST-tag and size-exclusion chromatography (see Chapter 2.1.4). The SNA1 PDZ domain was expressed at a lower level (18 - 20 mg/ml).

For the expression of the individual PDZ domains, a variety of growth parameters were altered to improve the yield of soluble fusion protein. In general, induction at lower cell densities ($E_{600} < 0.5$) usually resulted in larger yields of fusion protein in a soluble form. However, in some cases it was beneficial to grow the cells to a higher cell density ($E_{600} > 1.0$) and to induce for a shorter period of time with a reduced IPTG concentration (0.1 mM). These conditions often lead to lower yields of fusion protein, but prevented the proteolytical degradation.

2.1.4 PDZ Domain Purification

Cells were thawed and then lysed using a FRENCH®press (ThermoSpectronic, Rochester, USA), followed by centrifugation for 30 min at 75,800 g (25.50 rotor, Beckman, Munich, Germany). The supernatant was loaded onto a SP-sepharose column (Amersham Biosiences, Freiburg, Germany), where the GST-tagged proteins interacted with the

column. The column was first washed with PBS buffer pH 7.4 and the GST-tagged proteins were then eluted with 50 mM Tris buffer pH 8.0, 10 mM reduced glutathione. Fractions containing the GST-tagged protein of interest were pooled (Figure 2.1 (A)), then concentrated using a Vivaspin concentrator (Vivascience AG, Hannover, Germany) and used for the binding studies of membrane-bound peptide libraries incubations (see Chapter 0). For these incubations, the sample was concentrated to 1-5 mg/ml. It is not an advantage to use a protein solution with a higher concentration, because of the probability of protein precipitation. Once precipitated, the protein sample can not be turned back to solution. Additionally, the cleavage of the GST-tag by endogenous proteases could be observed at higher concentration.

For NMR measurements, the GST-tag of the AF6 PDZ domain had to be cleaved using PreScission™ protease. This protease is a genetically engineered fusion protein consisting of human rhinovirus 3C protease and GST [108], which was specifically designed to facilitate the removal of the protease by allowing simultaneous protease immobilization and cleavage of GST fusion proteins produced from the pGEX-6P vectors. PreScission™ protease specifically cleaves between the E and G residues of the recognition sequence LEVLFQ/GP [109].

After the first purification on the SP-sepharose column the fractions with the GST-tagged protein were pooled and dialyzed overnight against PreScisson™ cleavage buffer (50 mM Tris buffer pH 7.0, 150 mM NaCl, 1 mM EDTA). The GST cleavage was performed by adding 1 unit PreScission™ / µg protein for 4 h at 4°C. After the tag cleavage, we used a second SP-sepharose column to separate the GST-tag from the PDZ domain (Figure 2.1 (B)). The fraction containing the tag-free PDZ domain were then concentrated (Amicon Stirred Cell 8050, MY3 filter, Millipore, MA, USA) and further purified over a Superdex-75 column (Amersham Biosiences, Freiburg, Germany) in PBS buffer pH 7.4 to remove all impurities (Figure 2.1 (C)). The fractions containing the AF6 PDZ domain were pooled, concentrated to a 1-5 mg/ml solution and the buffer was changed to 20 mM phosphate buffer, 50 mM NaCl, Complete®, pH 7.0.

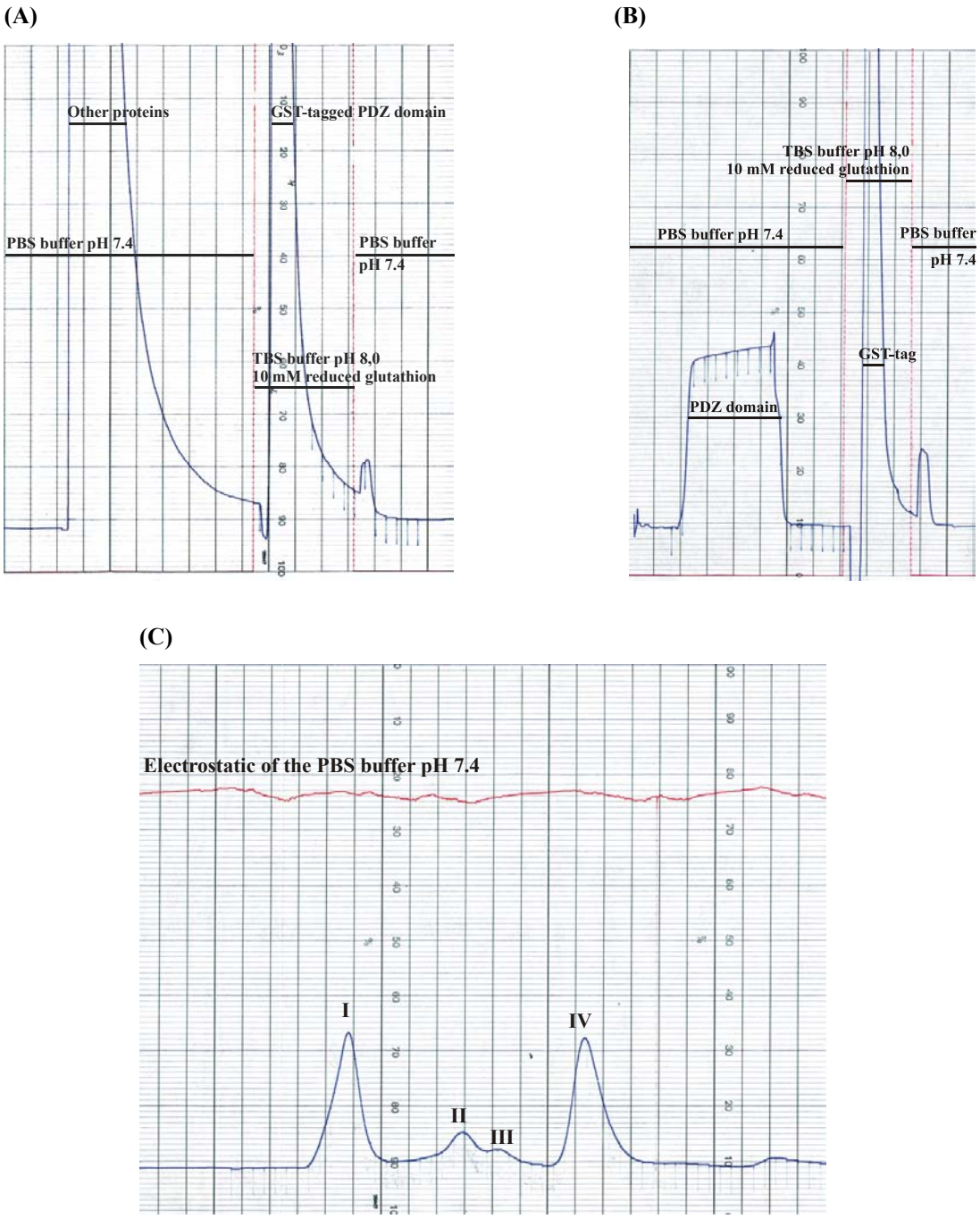

Figure 2.1 Chromatograms of the Different Protein Purification Steps.
(**A**) Chromatogram of the first SP-sepharose column used to extract the GST-tagged PDZ domain from cell lyses.
(**B**) Chromatogram of the second SP-sepharose column used to separate the PDZ domain from the GST-tag after GST cleavage with PreScission™ protease.
(**C**) Chromatogram of the final sample purification using Superdex-75 column showed the different peaks representing the PreScission^{TM} protease (I), the GST-tagged PDZ domain (II), the GST-tag (III) and the PDZ domain (IV).

2.1.5 PDZ Domain Analysis

2.1.5.1 Determination of Protein Concentration

All protein concentrations were determined by measuring light absorption at 280 nm, using Lambert-Beer's law ($E = \varepsilon\ c\ d$). For each PDZ domain the extinction coefficients (ε) at a wavelength of 280 nm (Table 2.2) were determined using ProtParam tool (www.expasy.org/tools/protparam.html) [110]. All protein solutions were diluted in water to allow a measurement of the extinction (E_{280}) between 0.5 and 1.0 (Ultraspec 3000, Amersham Biosiences, Freiburg, Germany).

Table 2.2 Characterization of the Expressed Proteins.

Protein	aa	MW [g/mol]	ε [M^{-1} cm^{-1}]
GST-AF6 PDZ	344	38354.4	43480
^{15}N-AF6 PDZ	118	11941.7	2560
GST-ERBIN PDZ	363	41364.7	44760
GST-SYNA1 PDZ	343	38957.0	43480

Footnotes: aa= amino acid, MW = molecular weight, ε = extinction coefficient at 280 nm.

2.1.5.2 Determination of Protein Purity

SDS-PAGE and mass spectrometry (MS) are standard methods to check the purity of the expressed proteins. SDS-PAGE [107] was mainly used to pursue the amount of protein through the expression and the purification process. The SDS-PAGE molecular weight marker (low range, Bio-Rad, Munich, Germany) was used as a reference. The proteins were visualized with Coomassie Brilliant Blue solution.

The SDS-PAGE of the ^{15}N-AF6 and the ERBIN PDZ domains (Figure 2.2) shows the expression and purification levels in minimal (A) and LB media (B). After IPTG induction, we clearly determine the protein expression observed as a protein lane at 11.9 kDa for the ^{15}N-AF6 (lane 3 in Figure 2.2 (A)) and a lane at 41.4 kDa for the ERBIN PDZ domain (lane 9 Figure 2.2 (B)). Moreover, after the SP-sepharose purification (lane 4 and 10 in Figure 2.2 (A, B)) the cleavage of the GST-tag (26 kDa) by endogenous proteases is detectable. The GST-tag could be removed during the purification with the Sephadex-75 column (lane 6 in Figure 2.2 (A)).

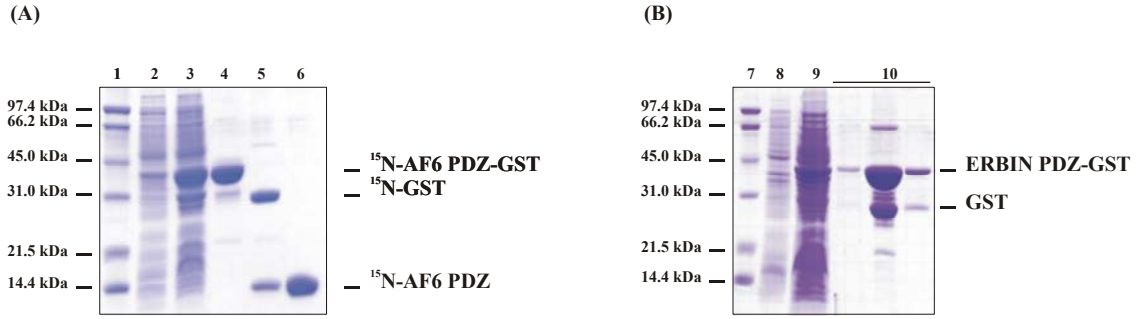

Figure 2.2 Characterisation of the Expression and the Purification of the AF6 and the ERBIN PDZ Domains.

(**A**) ^{15}N-AF6 PDZ domain expression in M9 media; (**B**). ERBIN PDZ domain expression in LB media. In both cases, the PDZ domains are expressed in the soluble fraction of the BL21 (DE3) cell lysate. After the purification steps, a cleavage of the GST-fusion protein by endogenous proteases is observed. Lane 1/7: molecular weight marker; Lane 2/8: cell extract without IPTG induction; lane 3/9: after addition of 1 mM IPTG; Lane 4/10: after the first SP-sepharose column purification; Lane 5: after GST-cleavage with PreScission™ protease and the second SP-sepharose column purification; Lane 6: after Superdex-75 column purification.

For the MALDI-TOF-MS analysis (Voyager-DE STR, Applied Biosystems, CA, USA), the protein solution was diluted by water to a final protein concentration of 1 mg/ml. The loss of one to ten amino acids (~ 1 kDa difference) resulting from protease activity during protein purification or tag-cleavage is sometimes detected by Matrix-Assisted Laser Desorption/Ionization Time of Flight (MALDI- TOF) mass spectrometry (MS). Figure 2.3 shows the mass spectra of the ^{15}N-labelled AF6 PDZ domain; the peak at 11934 kDa is in good accordance to the molecular weight of about 11941.7 kDa, assuming 100% ^{15}N-labelling.

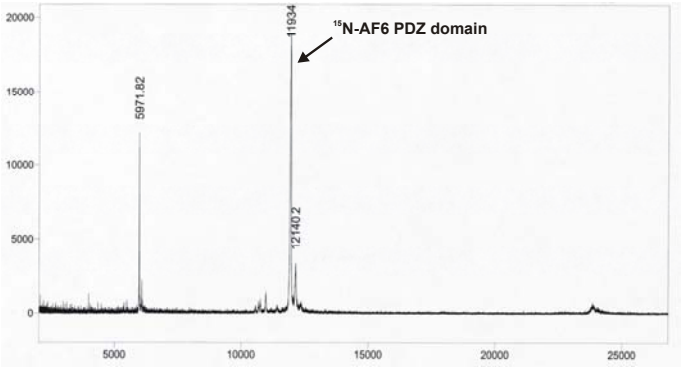

Figure 2.3 MALDI-TOF Mass Spectra of the ^{15}N-AF6 PDZ Domain.

2.1.5.3 Determination of Protein Folding

1D-^{1}H-NMR and circular dichroism (CD) measurements were applied to investigate the folding state of the samples.

For circular dichroism (CD) measurements, the sample concentration should ideally be chosen to give an Emax ~ 0.8, which occurs at protein concentration in the range of 10 - 100 µM. Emax includes absorption from all species in the measured solution including solvent and buffer salts. For this reason, it is important that the chloride concentration is kept at a minimum by dissolving the sample in water. Typically, a wavelength range of 190 - 260 nm (scanning speed 50 nm/min) was used for the measurement. Figure 2.4 (A) shows as an example the CD spectra of the AF6 PDZ domain with a minimum observed between 205 and 220 nm. This represents an overlay of structural contributions from α-helices and anti-parallel β-strands, as present in the PDZ domain structure.

The ^{1}H-NMR measurement was performed with the AF6 PDZ domain (1.24 mg/ml, 20 mM phosphate, 50 mM NaCl, Complete®, pH 7.0) at 300 K. 10% d_6DMSO was added to the protein solution and used as lock signal. The spectrometer DRX600 was operated using the program ICON-NMR on top of XWIN-NMR (Bruker, Karlsruhe, Germany).

(A) **(B)**

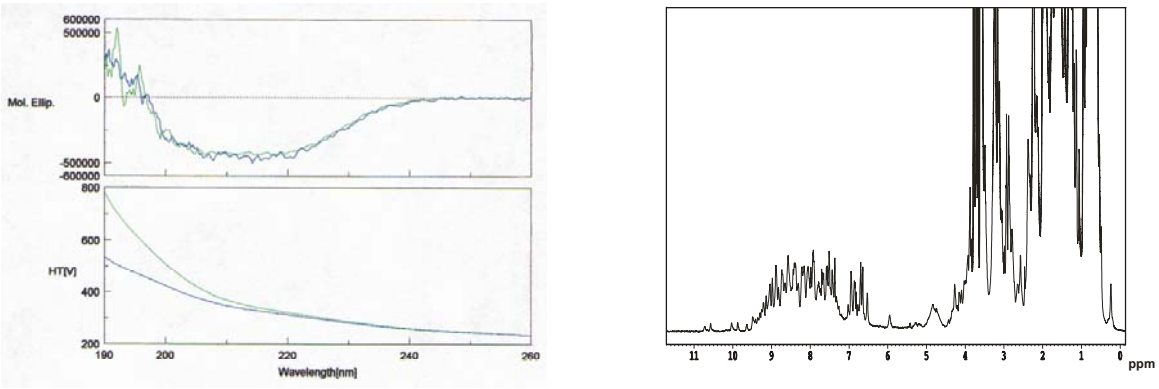

Figure 2.4 Determination of the Folding State of the AF6 PDZ Domain by CD and NMR.
(**A**) CD Spectroscopy: AF6 PDZ domain measured in a wavelength range of 190 - 260 nm. The measurement was performed at to sample concentration: blue line at 100 µM (thickness of the cuvette: d = 0.1 cm) and green line at 26.8 µM (d = 0.095 cm).
(**B**) 1D ^{1}H-NMR spectroscopy: chemical shift dispersion of the AF6 PDZ domain.

The 1D-^{1}H-NMR spectrum of the AF6 PDZ domain is shown in Figure 2.4 (B). The folding of the PDZ domain can be deduced through the pattern of the chemical shifts of the backbone amide protons: in an unfolded protein the chemical shifts of the amide protons are all determined by the same contacts with the solvent, resulting in values around 8.3 ppm. This spectral area is a region characteristic for backbone amides in random-coil conformation and, thereby, an excellent indicator for a disordered protein. On the other hand, signal dispersion beyond 8.7 ppm is a proof for a folded state of a protein. Because the varying shielding effects in a protein structure, the resonances are distributed over a wide range of frequencies in a spectrum of a folded protein (Figure 2.4 (B)). Following the same argument, large signal dispersions in the aliphatic region of the spectrum (1.0 to -1.0 ppm) *versus* a steep flank of a dominant peak (~1 ppm) distinguish a structured from an unfolded protein.

2.2 SPOT Synthesis

2.2.1 Synthesis of the Fmoc-Amino Acid 3-Bromopropyl Esters

Pyridine (1.6 ml, 10 mmol) was added to a stirred mixture of the respective Fmoc-amino acid (10 mmol) in dichloromethane (100 ml), followed by the addition of cyanurfluoride (1.5 g, 11 mmol). The mixture was stirred for 3 h at room temperature and then poured on ice water. Subsequently, the organic layer was separated, washed with a saturated solution of salt (3x), dried ($MgSO_4$), filtered, dried in vacuum and coevaporated twice with toluene. The resulting solid was redissolved in dichloromethane (50 ml) and mixed with 1-bromo-3-propanol (875 µl, 10 mmol) followed by addition of N-methylimidazole (795 µl, 10 mmol). After 1 h the reaction mixture was washed twice with citric acid (5%), with a saturated solution of salt (2x) and dried ($MgSO_4$). Evaporation usually resulted in an oil, which was coevaporated with diethyl ether to give a white solid substance. The crude products were purified by column chromatography performed on silica 60H (230 - 400 mesh), using diethyl ether/hexane as an eluent. The purity was checked by analytical RP-HPLC (Waters, Eschborn, Germany) on a Vydac C18 column (250 x 4.6 mm, Vydac, Hesparia, CA) using a linear gradient of 5 - 60% acetonitrile (0.05% trifluoroacetic acid

(TFA)) versus water (0.05% TFA) for 20 min at 1.2 ml/min flow rate. Absorbance was detected at 214 nm (A_{214}). Mass spectra were recorded on a quadrupole time-of-flight mass spectrometer type Q-TOF micro (Micromass, Manchester, UK) equipped with a nanoESI source. NMR-spectra were recorded at 300 K on Bruker DRX600 spectrometers in standard configuration and data were processed using XWIN NMR (Bruker, Karlsruhe, Germany).

2.2.2 Synthesis of Membrane-Bound Inverted Peptide Arrays

The peptides were synthesized on N-modified cellulose-amino-hydroxypropyl ether (N-CAPE) membrane [111] and prepared semi-automatically using a SPOT-robot (INTAVIS Bioanalytical Instruments AG, Cologne, Germany). Array design was performed using the in-house software LISA 1.571. The synthesis started with the spot definition by 9-fluorenyl-methoxycarbonyl-β-alanine-pentafluorophenyl ester (Fmoc-βAla-Opfp) using a standard protocol [112, 113], followed by the coupling of a 0.3 M solution of Fmoc-cysteine-(Trt)-OPfp in N-methyl-pyrolidone (NMP) (double coupling, 15 min each). After Fmoc-cleavage with 20% piperidine in N,N'-Dimethylacetamide (DMA), Mmt-mercaptopropionic acid was dissolved in dimethylformamide (DMF - 1 M solution) and activated with 1 equiv. HATU (N-[dimethylamino)-1H-1,2,3,-triazolo[4,5-b]-pyridin-1-ylmethylene]-N-methylmethanami-nium-hexafluorophosphate n-oxide) and 2 equiv. N-methylimidazole (NMI), and directly spotted on the membrane; this was repeated 15 min later. The membrane was washed with DMA (3 x 3 min) and dichlormethane (DCM) (3 x 3 min). The Mmt-group was removed using treatments of 10% (v/v) dichloroacetic acid and 0.5% (v/v) TFA in DCM for 1 x 5 min followed by 10% (v/v) dichloroacetic acid, 0.5% (v/v) TFA and 5% (v/v) triisobutylsilane (TIBS) in DCM for 3 x 5 min. The membrane was washed with DCM (1 x 3 min), EtOH (2 x 3 min), water (2 x 3 min), 10% aqueous solution of caesium carbonate to form the caesium thiolate (1 x 2 min), water (1 x 30 sec), EtOH (2 x 30 sec), diethyl ether (2 x 30 sec) and finally air-dried. Solutions of the Fmoc-amino acid 3-bromopropyl esters in DMF were spotted on the membrane (0.6 M solutions; 0.8 M solutions for C, H, N, Q and R; triple coupling, 15 min each). The Fmoc-group was removed from the spots and the sequences of the peptides were completed using the standard SPOT synthesis protocol [112, 113] and followed by a N-terminal tag with

β-alanine. For standard SPOT synthesis Fmoc-aa-OPfp were used with the following side-chain protections: E-, N-(O*t*Bu); S-, T-, Y-(*t*Bu); K-, W-(Boc); N-, Q-, H-(Trt); R-(Pbf) (Novabiochem, Laeufelfingen, Switzerland; Bachem, Budendorf, Switzerland). For thioether-cyclization all peptides were N-acylated with bromoacetic acid 2,4-dinitrophenyl ester in NMP (1 M), double coupling, 15 min each. The membrane was washed with DMA (3 x 3 min), DCM (3 x 3 min) and dried. Side-chain protection groups were removed using one treatment of 90% TFA and 3% TIBS in DCM for 30 min without shaking, washing steps (DCM 3 x 3 min, DMA 3 x 3 min, EtOH 3 x 3 min, diethyl ether 2 x 3 min) followed by 50% TFA and 3% TIBS in DCM for 2.5 hours without shaking. The membrane was washed with DCM (3 x 3 min), DMA (3 x 3 min), EtOH (2 x 3 min), 0.1 M phosphate buffer pH 7.4 (2 x 3 min) and water (2 x 3 min). The membrane-bound peptides were cyclized by incubating the membrane with a 5% aqueous solution of caesium carbonate for approximately 12 h. The membrane was washed five times with water and the membrane-bound cyclized peptides were inverted by ester-hydrolysis using one treatment with a saturated aqueous solution of lithium carbonate. Finally, the membrane was washed with water (3 x 3 min), 1% hydrochloric acid (1 x 3 min), water (2 x 3 min), EtOH (2 x 3 min), diethyl ether (2 x 3 min) and air-dried.

2.2.3 SPOT Synthesis and Chemical Characterization of the Peptides (11,12)

Synthesis of peptides was performed by manually pipetting 1 µl of each reagent solution onto the spots (spot area: 0.25 cm^2). Three arrays were generated, each consisting of eight identical spots. Cleavability of the prepared peptides and spot definition were achieved by spot-wise treatment of the N-CAPE membrane with a TBTU-activated 0.3 M solution of the Fmoc-Rink linker (Calbiochem-Novabiochem GmbH, Bad Soden, Germany) in DMF (1 equiv. Fmoc-Rink linker, 1 equiv. TBTU and 2 equiv. DIEA). After 15 min the procedure was repeated using a freshly prepared reagent solution. Synthesis of the peptides (**10**, Figure 3.5) by coupling Fmoc-cysteine-(Trt)-OPfp, Mmt-mercaptopropionic acid and standard SPOT synthesis was performed according to the protocol described above. Removal of the cysteine Trt-group was achieved by treatments of 2% (v/v) TFA in DCM for 1 x 5 min and 2% (v/v) TFA, 5% (v/v) TIBS in DCM for 3 x 5 min. The membrane was then washed with DCM (3 x 3 min) and air-dried. The three arrays were separated

cutting the membranes. For cyclization, each part was treated differently: (h-1) incubation with DMF for 3 h; (h-2) the membrane was washed with EtOH (2 x 3 min) and water (2 x 3 min), then incubated with a 5% aqueous solution of caesium carbonate for 1 min, washed with water (6 x 1 min), EtOH (1 x 3min), DMF (2 x 3 min) and incubated with DMF for 3 h; (h-3) the membrane was washed with DMF (2 x 3 min) and incubated with a 1 mM solution of caesium carbonate in DMF for 3 h. Subsequently, all membranes were washed with DMF (2 x 3min), DCM (3 x 3 min) and dried. The spots were punched out from each membrane and transferred to eppendorf tubes. Cleavage from the solid support and amino acid side-chain deprotection were performed simultaneously by adding a solution (300 µl) of 90% TFA, 5% water and 3% TIBS in DCM to each tube. The tubes were shaken for 3 h, the cellulose membranes were taken out and the resulting solution was concentrated *in vacuo*. Peptide precipitation was achieved by the addition of 100 µl diethyl ether and completed by centrifugation. After separation, the pellet was washed five times with diethyl ether and finally dried. The compounds were dissolved in 50% aqueous acetonitrile (200 µl) and the purity was checked by analytical HPLC (Waters, Eschborn, Germany) on a Vydac C 18 column. Cyclic compounds (**11**, Figure 3.5) were identified using ESI mass spectrometry:

Cyclic peptide (**11**, Figure 3.5), $C_{50}H_{79}N_{13}O_{17}S_2$

ESI-MS: mass calculated (monoisotopic): 1197.5158, mass found: 1197.6578 (found as m/z: 599.7657 $[M+2H]^{2+}$); HPLC: retention time: 12.01 min.

Linear peptide (**12**,Figure 3.5), $C_{50}H_{80}BrN_{13}O_{17}S_2$

ESI-MS: mass calculated (monoisotopic): 1277.4437, mass found: 1277.5972 (found as m/z: 639.7986 $[M+2H]^{2+}$); HPLC: retention time: 10.11 min.

2.2.4 Coupling Efficiency of Fmoc-amino acid 3-bromo-propyl esters (3).

Fifty-seven spots of the caesium thiolate (**4**, Figure 3.3) (spot area: 0.25 cm^2) on a N-CAPE membrane were treated with Fmoc-amino acid 3-bromo-propyl esters (Fmoc-aa-OPBr) (**3**, Figure 3.3) in DMF. Three spots were used for each 3-bromopropyl ester derivate (19 bromopropyl esters were tested, cysteine-adduct omitted). Each spot was treated with 1 µl of the respective Fmoc-aa-OPBr (**3**, Figure 3.3) solution. We used solutions with a concentration of 0.6 M (2x coupling) and of 0.8 M (3x coupling).

Coupling yields were determined by measuring the UV-absorbance of the released Fmoc-group [111] and are given as the mean loading of one spot (c[nmol/spot] = OD_{301} / ε * 1000, with ε = 8.1).

2.2.5 Binding Studies of Cellulose-Bound Peptides

All libraries were generated by a semi-automated SPOT-robot [112, 113] (Abimed, Langenfeld, Germany; Software LISA, in-house software) using the novel strategy of 'inverted peptides' (see Chapter 2.2.2).

These libraries were pre-washed once with EtOH (1 x 10 min), with Tris-buffered saline (TBS buffer) pH 8.0 (3 x 10 min) and then blocked for 4 h with blocking buffer (blocking reagent (Sigma-Genosys, Cambridge, USA) in TBS buffer pH 8.0, containing 5% sucrose). The membranes were incubated with the PDZ-GST (10 µg/ml) in blocking buffer overnight at 4°C, washed with TBS buffer pH 8.0 (3 x 10 min), followed by a second incubation with anti-GST antibody conjugated with the horse radish peroxidase (-HRP) (Amersham Biosiences, Freiburg, Germany) (1 µg/ml) in blocking buffer for 3 h at room temperature. To remove excess antibody the membrane was washed with TBS buffer pH 8.0 (3 x 10 min). A SuperSignal West Pico chemiluminescence substrate (Pierce, Rockford, USA) was applied for detection using a LumiImager™ (Boehringer Mannheim GmbH, Mannheim, Germany). The signal intensities were recorded as Boehringer Light Units (BLU) using the LumiAnalyst™ software.

2.2.6 Peptide Synthesis on Resin

The peptides used for the binding studies were automatically synthesized (AMS 422, Abimed, Langenfeld, Germany) according to the Fmoc standard protocol [112]. The generation of peptides with a C-terminal carboxyl group and with C-terminal amide was performed using TentaGel S PHB-aa-Fmoc and TentaGel S Ram resin, respectively. For standard synthesis Fmoc-aa-OH were used with the following side-chain protections: E-, D-(O*t*Bu); S-, T-, Y-(*t*Bu); K-, W-(Boc); N-, Q-, H-(Trt); R-(Pbf) (Novabiochem, Laeufelfingen, Switzerland; Bachem, Budendorf, Switzerland).

Each cycle consisted in the N-terminal Fmoc cleavage by 20% piperidine (2 x 15 min), followed by DMA washing steps and ending with the coupling of the next amino acid.

Each amino acid is activated by 1 equiv. PyBOP ([benzotriazol-1-yloxy]tripyrrolidinophos-phonium hexafluorophosphate) and 2 equiv. NMI (2 x 15 min). Before the following Fmoc-cleavage step, the resin was washed 6 times with DMA.

The cleavage of the side-chain protection groups and of the whole peptide was processed using a cleavage-solution of TFA, phenol, water, methylphenylsulfide and ethandithiole in a ration of 20:1.5:1:1:0.5 for 3 h. Thereafter, the peptides were precipitated with ice-cold *trt*-butylmethyl ether and centrifuged 2 min at 3200 rpm. The precipitate was washed three times with diethyl ether and centrifuged again (2 x 2 min at 3200 rpm, 1 x 6 min at 5800 rpm).

The crude peptides were purified using preparative RP-HPLC (Waters, Eschborn, Germany) on a Vydac C 18 column, analyzed by analytical RP-HPLC and ESI-MS. The purified peptides were finally lyophilized. All peptides exhibited finally a free N-terminal amino group.

2.2.7 Swiss-Prot Database Accession Number of the Used Peptides and Proteins

PDZ domains deduced from the following proteins were used in this work: AF6 [Swiss-Prot (SP): P55196], ERBIN [SP: Q9NR18] and SNA1 [SP: Q61234].

Based on the peptide library containing 6223 non-redundant human C-terminal sequences, the following C-terminal peptide sequences or proteins were analyzed in detail as PDZ domain ligands: A1AD [P25100], AG22 [SP: P50052], APC [P25054], ATB1 [SP: P20020], ATB2 [Q01814], ARVC [SP: O00192], AVR2 [P27037], BCR [SP: P11274], CIK4 [SP: P22459], CIK5 [SP: P22460], CIN4 [SP: P35499], CTNB [P35222], CYG4 [SP: P33402], EPA7 [Q15375], EPB2 [P29323-2], ERB2 [P04626], FXI1 [Q12951], MCM7 [P33993], NME2 [Q13244], PTPZ [P23471], REL [Q04864], RGSC [O14924], RHM1 [SP: P25800] and TAT [P03409].

2.3 Biophysical Methods

2.3.1 Affinity Measurements

Binding constants (K_d) were measured using the BIACORE® X system (Uppsala, Sweden). PDZ-GST was immobilized on a CM5 chip using the GST coupling procedure, according to the supplier's instructions. The amount of covalently coupled GST-tagged protein in flowcell 2 corresponded to a signal increase of approximately 500 resonance units (RU). An appropriate amount of GST was coupled to flowcell 1 on the same chip as a reference.

All binding experiments were performed at 20°C with a flow rate of 5 µl/min (injection volume 10 µl). Peptides were used at various concentrations between 1 µM and 1 mM in HBS buffer (10 mM Hepes with 0.15 M NaCl, 3.4 mM EDTA, and 0.005% surfactant P20, pH 7.4). Complete regeneration was obtained after dissociation without regeneration buffer. Transformation of data and analysis were performed with the BIA-evaluation software, version 3.0. The control sensorgram (flowcell 1) was subtracted from the sensorgrams obtained with flowcell 2. The steady-state values of the binding equilibrium were plotted versus the different peptide concentrations and fitted using the implemented steady-state evaluation, resulting in the K_d for the PDZ domain/peptide complexes.

2.3.2 NMR Spectroscopy

2.3.2.1 Backbone Assignment of the AF6 PDZ Domain

The NMR samples typically contained 1.3 mM ^{15}N-^{13}C-labeled AF6 PDZ domain, 20 mM phosphate buffer (pH 7.0), 50 mM NaCl and Complete® protease inhibitor and 10% D_2O. NMR spectra were acquired at 300 K on Bruker DRX600 spectrometers in standard configuration using an inverse triple resonance probe with Z-gradient. Data were processed using XWIN-NMR and analyzed applying Sparky (University of California, San Francisco, USA). Backbone resonance assignment was carried out based on a complete series of side chain selective ^{1}H-^{15}N-HSQC-experiments [114-117] supported by CBCA(CO)NNH/ CBCANNH and HA(CO)NNH/HANNH experiments [118, 119].

2.3.2.2 Chemical Shift Experiments

Aliquots of peptide were added to the ^{15}N-labeled AF6 PDZ domain until ^{1}H-^{15}N-HSQC peaks for the free protein had disappeared and no further change of intensity for the peaks of the complexed protein were observable. Total shifts perturbations ($\Delta\delta_{total}$) of ^{1}H and ^{15}N resonances were calculated according to $\Delta\delta_{total} = [(\Delta^{1}H)^2 + (\Delta^{15}N\ 0.2)^2]^{½}$.

Dissociation constants were obtained by NMR titration of a 0.1 mM ^{15}N-labeled AF6 PDZ domain sample with increasing peptide concentrations ranging from 30 µM to 10 mM. The K_d value for the PDZ domain/peptide complex was determined as reported previously in [120].

2.3.3 Molecular Modeling and Molecular Dynamics Simulation

2.3.3.1 The AF6 PDZ Domain/Peptide Complex

The AF6 PDZ domain model has been built by homology modeling based on the SNA1 PDZ domain (PDB: 2PDZ) [34] template with ~40 % sequence identity. All model components were assembled with the biopolymer module of the SYBYL program package version 6.6 (TRIPOS Inc. St Louis, MO, USA). Molecular dynamics (MD) simulations were performed at 300 K for 500 ps, using the AMBER 4.1. The geometrical quality of the resulting model was checked using the program ProCheck [121].

2.3.3.2 The ERBIN PDZ Domain/Peptide Complex

The ERBIN PDZ domain structure in complex with the C-terminal peptide of angiotensin type II receptor (peptide sequence: METFVS$_{COOH}$) was built by homology modeling based on the structure 1MFG as a template [122]. The ligand was placed in the binding pocket based on structural information obtained from all PDZ domain/ligand complexes available in the PDB database. Adding counter ions neutralized the charge of the whole system, which was finally embedded into an octahedral water box for periodic simulations. This model complex structure was then subjected to MD simulations using AMBER 7 [123] in conjunction with the force field of Cornell *et al.* [124]. After a short energy minimization in vacuum for 1000 steps, a harmonic potential with a force constant of 30 kcal mol^{-1} Å^2 was applied to restrain all atoms belonging to the protein. During the equilibration phase the system was gradually heated from 100 K to 300 K (25 ps) under otherwise constant

conditions and with the protein kept constraint. During repeated short minimization (600 steps, each) the restraints on the protein were gradually (in 5 kcal mol-^{1} Å^{2} steps) relieved until a completely unrestrained system was obtained. The complex was then subjected to unrestrained MD for another 20 ps followed by simulations for 1000 ps under conditions of constant temperature (300 K) and pressure (1013 mbar). Trajectory frames were collected every 1 ps for detailed analysis. The trajectory was analyzed using the tools ptraj and carnal. We selected the structure with the lowest RMSD to the average structure between 700 and 900 ps for further analysis.
All figures (ribbon models, overlays and surface models) were generated using MolMol 2K.2 [125].

2.3.4 Prediction of Dissociation Constants

These experiments were done in cooperation with Urs Wiedemann (Group of Prof. Dr. H. Oschkinat, FMP, Berlin, Germany).

2.3.4.1 Calibration of BLU Values

The BLU values obtained from the incubation of membrane-bound peptide libraries can be related to the underlying dissociation constants by the following two equations:
First, the measured *BLU* values are directly proportional to the amount of enzyme bound to the membrane (E_{bound}) (equation (1)), which is in turn directly proportional to the amount of bound PDZ domain (D_{bound}) (equation (1')):

$$BLU = \varepsilon E_{bound} + \alpha \quad (1)$$

$$E_{bound} = \kappa D_{bound} \quad (1')$$

$$BLU = \lambda D_{bound} + \alpha \quad (1'')$$

We assume a large excess of enzyme substrate and merge the proportionality factor ε (summarizing the enzyme-specific Michaelis Menten constant K_M, the incubation time and the size of the detection area) and the proportionality factor κ for potential indirect detection systems (e.g. the K_d of enzyme-labeled secondary antibodies) into a single proportionality constant λ. The background correction is estimated as the offset parameter α (equation (1'')).

Secondly, the amount of bound PDZ domain D_{bound} is related to the dissociation constant via the law of mass action (equation (2)) which can be transformed to its inverse (equation (2')) assuming a 1:1 binding model:

$$K_d = [\ (D_{total} - D_{bound})\ (P_{total} - D_{bound})\ ]\ D_{bound}^{-1} \quad (2)$$

$$D_{bound} = ½\ (K_d + D_{total} + P_{total} - [\ (-K_d - D_{total} - P_{total})^2 - 4\ D_{total}\ P_{total}\ ]^{½}\ ) \quad (2')$$

The total amount of applied PDZ domain D_{total} and the corresponding amount of available membrane bound peptides P_{total} are assumed to be constant parameters for all spots on one membrane subjected to a single incubation.

Equations (1) and (2) together yield the model relating K_d to *BLU*, using the log-transformed *BLU* values to account for their approximate log-normal distribution:

$$\ln[BLU - \alpha] = \ln[\lambda\ ½\ (K_d + D_{total} + P_{total} - [\ (-K_d - D_{total} - P_{total})^2 - 4\ D_{total}\ P_{total}\ ]^{½}\ )] \quad (3)$$

The profile libraries were accompanied by 5 replicate calibration spots for each peptide where we had determined the K_d by SPR measurements (calibration peptides). After the maximum likelihood estimation of the offset factor α using *logtrans* [126] the parameters λ, D_{total} and P_{total} were estimated by non-linear least-squares fitting of the model (3) to the BLU values of the calibration peptide. Using these parameter estimates we modeled the K_d values corresponding to the measured BLU for the profile library sequences.

2.3.4.2 Analysis of Variance (ANOVA) Model of Specificity

The profile libraries for the AF6 and ERBIN PDZ domains represent complete factorial designs with the calibrated BLU values (expressed as log-transformed K_d modeled from the measured *BLU*) as the response. Using the sequence of the 4 C-terminal positions, 4-way fixed effects [127] ANOVA models without statistical interaction terms were built (Equation 4, assuming no cooperativity between individual ligand residue contributions α_{PosX}):

$$y_i = \alpha_{Pos\text{-}3} + \alpha_{Pos\text{-}2} + \alpha_{Pos\text{-}1} + \alpha_{Pos0} \quad + \quad \varepsilon_i \quad (4)$$

In addition, 4-way fixed effects ANOVA models with pair-wise statistical interaction terms were built (Equation 5):

$$y_i = \alpha_{Pos\text{-}3} + \alpha_{Pos\text{-}2} + \alpha_{Pos\text{-}1} + \alpha_{Pos0} + \alpha_{Pos\text{-}3}{:}\alpha_{Pos\text{-}2} + \alpha_{Pos\text{-}3}{:}\alpha_{Pos\text{-}1} + \alpha_{Pos\text{-}3}{:}\alpha_{Pos0} + \alpha_{Pos\text{-}2}{:}\alpha_{Pos\text{-}1} + \alpha_{Pos\text{-}2}{:}\alpha_{Pos0} + \alpha_{Pos\text{-}1}{:}\alpha_{Pos0} + \varepsilon_i \quad (5)$$

where y is the response vector, with y_i being the ith observation and ε_i its corresponding residual variation. The vector of fixed effects for the amino acid types in ligand position X is α_{PosX} and the interaction term for cooperativity between position X and Y is $\alpha_{PosX}{:}\alpha_{PosY}$. The $\alpha_{PosX}{:}\alpha_{PosY}$ interaction term can also be read as the non-additive contribution of having amino acid type Xaa_1 in position X and Xaa_2 in position Y together in the same peptide.

In order to include also substitutional analyses, we used mixed effects [128] ANOVA models which allow to compensate for between-experiment errors (Equation 6):

$$y_{ni} = \alpha_{Pos\text{-}3} + \alpha_{Pos\text{-}2} + \alpha_{Pos\text{-}1} + \alpha_{Pos0} + \zeta_n + \varepsilon_{ni} \quad (6)$$

Mixed effects models accomplish this by decomposing the variation into random effects ζ at the level of the individual experiments (n experiments grouping the $n \times i$ observations) and the within-experiment error ε. The parameters of the mixed effects models were estimated by restricted maximum likelihood.

The full models above were validated by comparing the predicted K_d (based on the sequence of the four C-terminal residues) with the experimentally determined K_d values (validated R^2). In addition, we removed the data for those peptides with K_d values measured by SPR from the ANOVA data sets and built cross-validated models based on these reduced data sets. We calculated then the Q^2 (cross-validated R^2) by evaluating the fit of the predictions from the cross-validated models to the experimentally determined K_d values of those peptides that were left out before.

Statistical analysis was performed using the program R version 1.8.1 (http://www.R-project.org) [129].

2.4 Biological Methods

2.4.1 Pull-down and Coimmunoprecipitation Assays

These experiments were done in cooperation with Dr. Gerald Radziwill (Group of Prof. Dr. K. Moelling, Institute of Medical Virology, Zurich, Switzerland).

The full-length breakpoint cluster region (BCR_{WT}) and the mutant BCR_{V1271A} were described elsewhere [67]. Pull-down and immunoprecipitation of BCR and influenza hemaglutinin-epitope (HA)-tagged BCR were performed with rabbit polyclonal anti-BCR C20 (Santa Cruz Biotechnology) for BCR and Mab 12CA5 (Roche diagnostics) for HA-BCR. The gluthation-S-tranferase (GST)-tagged ERBIN PDZ and ERBIN were analyzed with anti-GST (Amersham Pharmacia Biotech) for GST-ERBIN and anti-ERBIN (in-house production, Institute of Medical Virology, Zurich, Switzerland) for ERBIN, respectively.

Human embryonic kidney 293 (HEK293) and human gastric adenocarcinoma (MKN7) cells were cultured at 37°C and with 5% CO_2 in Dulbecco's modified Eagle medium containing 10% fetal calf serum, 100 U penicillin/ml and 100 µg of streptomycin/ml. Transfection performed based on standard calcium-phosphate methods, pull-down and coimmunoprecipitaion were performed as described in [67].

2.4.2 Immunolocalization

These experiments were done in cooperation with Viola Weber (Group of Prof. Dr. W. Rosenthal, FMP, Berlin, Germany).

For immunostaining procedures, HEK292 cells were used and then transfected with the respective DNA constructs using FuGENE (Roche Diagnostic, Mannheim, Germany) (for details see [130]). The DNA-construct of the angiotensin type II receptor (AG22) C-terminus (11mer) is fused to the gene of the green fluorescence protein (GFP-$AG22_{Cterm}$), and could therefore be directly visualized. HA-tagged AG22 (whole receptor) and FLAG-tagged ERBIN (full-length) were visualized with polyclonal rabbit anti-HA (Roche Diagnostic, Mannheim, Germany)/anti-rabbit-cy3 (ABCAM, Cambridge, UK) for AG22 and monoclonal mouse anti-FLAG (ABCAM, Cambridge, UK)/anti-mouse-cy5 (ABCAM, Cambridge, UK) for ERBIN. Images were obtained using a Zeiss LCM 510meta microscope (100x1.3, oil).

Setting up the cells: Three autoclave coverslips were placed in one tissue culture dish. 10 ml of culture media was added to each dish. After splitting the cells, an appropriate amount of cells was added to each dish. The cells needed some time to adhere on coverslips, therefore the cells were keep at least 24 h after the transfection at 37°C. The

number of cells added to each dish depended on the individual cell growth rate. However, at the time of fixation cells should be ~ 40-50 % confluent. To be able to get good images under the microscope single isolated cells were needed. Culture media were drawed off and the dishes were washed 2x with warm PBS. The coverslips were transferred to 6-well culture dishes/one coverslip per well.

Fixation: 30µl solution of 100 mM cacodylate, 100 mM sucrose, 10% paraformaldehyde (PFA), pH 7.6 was added for 15 min at RT and thereafter washed 2x with PBS for 5 min. For PFA quenching, the coverslips were incubated with 50 mM ammonium acetate in PBS for 20 min and thereafter washed with PBS (2x).

Permeabilization: 30 µl of 1 % Triton in PBS were added for 10 min at RT and washed 2x with PBS for 5 min.

Antibody Incubation: The primary antibody was diluted in PBS as recommended by manufacturer specification data sheet. 30µl of the primary antibody in PBS were added and incubated in a humidified chamber for 45 min at 37°C. Subsequently, they were washed with PBS for 10 min (3x). Secondary antibody was also diluted in PBS according to the recommended manufacturer specification data sheet. 30 µl of each antibody were added per well and incubated in humidified chamber for 45 min at 37°C. Finally, they were washed with PBS for 10 min (3x).

Mounting: The coverslips were washed with tridest. water and dried at 37°C for ~ 45 min. 10-15 µl mounting media were used per coverslip (mounting media: 50-60 % glycerol, 2.5 % 1,4 diazobicyclo(2,2,2)-octane in PBS). The mountant was applied on slide and the coverslip with cells facing mountant on top were added and let it dry over night at 4°C.