Chapter 3

Material and Methods

3.1 Material

If not specified otherwise chemicals were purchased in the highest available quality from Sigma, Merck, Fluka, Roche or Serva.

3.1.1 Bacterial strains

- *E. coli* BL21(DE3) for protein expression (Novagen, Darmstadt, Germany)

3.1.2 Expression vectors

- modified pET9d vector (Novagen, Darmstadt, Germany) with kanamycin resistance for the production of fusion proteins with an N-terminal His-tag (MKH_6PMG), obtained from G. Stier
- modified pET24d vectors (Novagen, Darmstadt, Germany) with kanamycin resistance for the production of N-His-tagged (MKH_6PMG) and N-His_6-GST tagged fusion proteins with TEV (tobacco etch virus) protease cleavage site, obtained from G. Stier

3.1.3 Buffers and solutions

All buffers and solutions were sterile filtered prior to usage. All specifications refer to 1 litre solution.

- **Lysis buffer:** 50 mM Na phosphate pH 8.0, 10 mM imidazole pH 8.0, 150 mM NaCl, β-mercaptoethanol in 1:10.000 dilution
- **Lysis buffer A:** lysis buffer, 0.1% (w/v) Triton X-100

- **Washing buffer 1:** as lysis buffer, but 1 M NaCl
- **Washing buffer 2:** as lysis buffer, but 20 mM imidazole pH 8.0
- **Elution buffer:** as lysis buffer, but 300 mM imidazole pH 8.0
- **NMR buffer:** 20 mM Na phosphate, 30 mM NaCl, 0.03% (w/v) NaN_3.

  The pH of the NMR buffer varied from 5.8 to 7.2 depending on the pI of the dissolved protein.
- **Trace elements (100x stock solution):** 5 g EDTA, 0.833 g $FeCl_3\times$ 6 H_2O, 84 mg $ZnCl_2$, 13 mg $CuCl_2\times2$ H_2O, 10 mg $CoCl_2\times6$ H_2O, 10 mg H_3BO_3, 1.6 g $MnCl_2\times6$ H_2O, pH 7.5

  Prior to the addition of the remaining salts, EDTA must be dissolved in water and the pH adjusted to 7.5.

3.1.4 Cell culture media

- **LB-medium (1 litre):** 10 g bactotryptose, 5 g yeast extract, 10 g NaCl, pH 7.4, 1 ml 25 mg/ml kanamycin, autoclaved
- **LB-agar plates:** 1.5% (w/v) agar in autoclaved LB-medium, 25 μg/ml kanamycin
- **M9 medium (1 litre 10x stock solution):** 60 g Na_2HPO_4, 30 g KH_2PO_4, 5 g NaCl, 5 g NH_4Cl ($^{15}NH_4Cl$ for ^{15}N-labelled M9 medium), autoclaved
- **minimal medium (1 litre):** 100 ml 10x M9 medium, 10 ml 100x trace elements, 20 ml 20% glucose (10 ml 20% $^{13}C_6$-glucose for ^{13}C-labelled minimal medium), 1 ml 1M $MgSO_4$, 0.3 ml 1M $CaCl_2$, 1 ml 1 mg/ml biotin, 1 ml 1 mg/ml thiamin, H_2O ad 1 litre, 1 ml 25 mg/ml kanamycin, sterile filtered

3.1.5 Recombinant proteins

Additional N-terminal residues resulting from purification tags (MKH_6PMG) or TEV protease cleavage (GAM or GAMG) are separated from the given sequences. Otherwise the sequences are given in blocks of 10 residues such that the last residue in a block corresponds to a residue number which is a multiple of 10.

- unlabelled, ^{13}C- and/or ^{15}N-labelled tandem *Sc.* Prp40 WW domains (Prp40 1–75):
  * N-His-tagged:
  `MKHHHHHHPMG MSIWKEAKDA SGRIYYYNTL TKKSTWEKPK ELISQEELLL RENGWKAAKT ADGKVYYYNP TTRETSWTIP AFEKK`

* Non-tagged:
GAM MSIWKEAKDA SGRIYYYNTL TKKSTWEKPK ELISQEELLL RENGWKAAKT ADGKVYYYNP TTRETSWTIP AFEKK

- unlabelled and ^{15}N-labelled *Sc.* Prp40 WW2 domain (Prp40 42–75):
GAME RENGWKAAKT ADGKVYYYNP TTRETSWTIP AFEKK

- unlabelled, ^{13}C- and/or ^{15}N-labelled *Sc.* Prp40 FF1 domain (Prp40 134–189) provided by Dr. A. Gasch:
GAM EAEKEFI TMLKENQVDS TWSFSRIISE LGTRDPRYWM VDDDPLWKKE MFEKYLSNR

- ^{15}N-labelled *Sc.* Prp40 FF2 domain (Prp40 200–260) provided by Dr. A. Gasch:
GAM E TSKFKEAFQK MLQNNSHIKY YTRWPTAKRL IADEPIYKHS VVNEKTKRQT FQDYIDTLID

- ^{15}N-labelled *Sc.* Prp40 tandem FF1-2 domains (Prp40 134–260) provided by Dr. A. Gasch:
GAM EAEKEFI TMLKENQVDS TWSFSRIISE LGTRDPRYWM VDDDPLWKKE MFEKYLSNRS ADQLLKEHNE TSKFKEAFQK MLQNNSHIKY YTRWPTAKRL IADEPIYKHS VVNEKTKRQT FQDYIDTLID

- ^{15}N-labelled *Sc.* Prp40 FF6 domain constructs (Prp40 488–552 and 465–552, respectively) provided by Dr. A. Gasch:
* Prp40 488–552:
GAMG TLI DQNFEWNDAD SDEITKQNIE KVLENDRKFD KVDKEDISLI VDGLIKQRNE KIQQKLQNER RI

* Prp40 465–552:
GAM DLFLDF VDEQRMYIFA QRSIAQQTLI DQNFEWNDAD SDEITKQNIE KVLENDRKFD KVDKEDISLI VDGLIKQRNE KIQQKLQNER RI

- ^{15}N-labelled *Sc.* Ypr152 FF domain (Ypr152 212–266) provided by Dr. A. Gasch:
GAM IDERNIFFE LFDRYKLDKF STWSLQSKKI ENDPDFYKIR DDTVRESLFE EWCGER

- unlabelled, ^{13}C- and/or ^{15}N-labelled *Sc.* Clf1 TPR1 repeat (Clf1 31–64) provided by Dr. A. Gasch:
GAM GSTNIDILDL EELREYQRRK RTEYEGYLKR NRLD

- ^{15}N-labelled *Sc.* extended Clf1 TPR1 repeat (Clf1 1–64) provided by Dr. A. Gasch:
  `GAM MDTLEPTAVD THVSAEQILR DVYKKGQKAR GSTNIDILDL EELREYQRRK RTEYEGYLKR NRLD`

3.1.6 Synthetic peptides

- *Sc.* Prp8 6–14: `PPPPGFEED`
- *Sc.* Prp8 24–32: `PPPPGYEIE`
- *Sc.* Prp8 53–60: `PPPPSNFE`
- *Sc.* BBP 94–102: `PSPPPVYDA`
- *Sc.* CTD repeat: `SYSPTSPSYSPTSPS`
- *Sc.* phosphorylated CTD repeat: `SYpSPTpSPS`
- *Hs.* Formin 881–888: `PPLIPPPP`
- *Hs.* 3BP-10: `APTMPPPLPP`

All synthetic peptides were purchased unlabelled from MWG-Biotech AG (Ebersberg, Germany).

3.2 Biochemical methods

To produce the recombinant proteins used for NMR in this Thesis, the nucleotide sequence encoding the protein of interest is amplified by polymerase chain reaction (PCR) using genomic DNA as template. The gene fragment is inserted into a suitable circular DNA molecule, an expression vector, which can be introduced into competent host cells by transformation. In the pET (premier *E. coli* expression) system, the target gene is transcribed by the highly efficient T7 RNA polymerase (Fig. 3.1), which itself is expressed under stringent control of an inducible *lac*-promotor (DE3 fragment in the *E. coli* genome in Fig. 3.1) (Studier *et al.*, 1990). In the absence of inducer molecules, such as IPTG for *lac*-promotors, expression of T7 RNA polymerase and hence the target gene is repressed, since the promotor is blocked by the *lac*-repressor. Addition of IPTG releases the *lac*-repressor, so that T7 RNA polymerase and therefore also the target gene can be expressed. Some pET vectors also carry their own *lac*-repressor gene to ensure that enough repressor is made to bind to all available operator sites. Moreover, expression vectors contain antibiotic resistance genes to enable selection for plasmid containing host cells by growing the transformed cells on medium containing the respective antibiotics (*e.g.* kanamycin). To produce uniformly ^{13}C and/or ^{15}N-labelled recombinant protein cells are grown on minimal medium containing $^{13}C_6$-glucose and $^{15}NH_4Cl$ as sole sources of carbon and nitrogen, respectively.

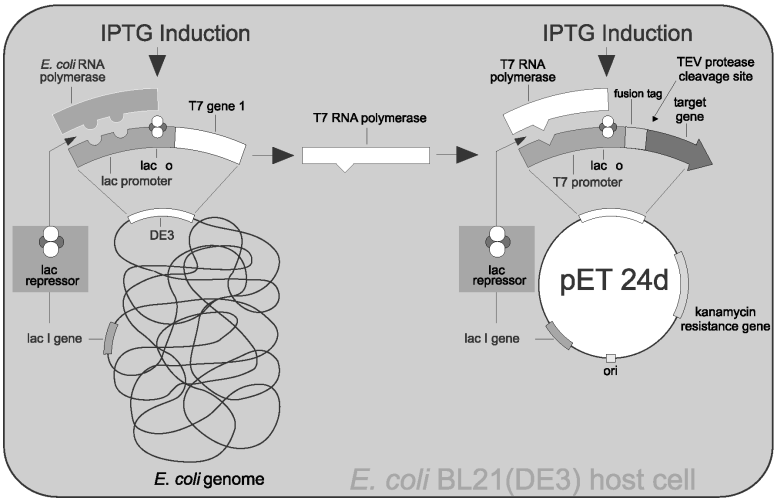

Figure 3.1: Schematic representation of the *E. coli* BL21(DE3) expression system showing a modified pET vector as used in this Thesis. The Figure is adopted from a Novagen (Darmstadt, Germany) illustration.

Furthermore it is convenient to fuse the target gene with purification tags in the expression vector so that the fusion protein can be separated from unwanted proteins by affinity chromatography using the corresponding resins. In this Thesis, N-terminal histidine-tags were used, which bind specifically to Ni-resins. A TEV (tobacco etch virus) protease cleavage site was cloned between the tag and the sequence of the target protein to cleave off the tag. After cleavage the pure target protein was obtained in the flow-through of another Ni-affinity chromatography step using His-tagged TEV protease (provided by G. Stier or the protein expression unit at EMBL Heidelberg) (see Fig. 3.2). All plasmids used in this Thesis were sequenced (at SEQLAB, Göttingen, Germany) and provided by G. Stier, Dr. M. J. Macias or Dr. A. Gasch.

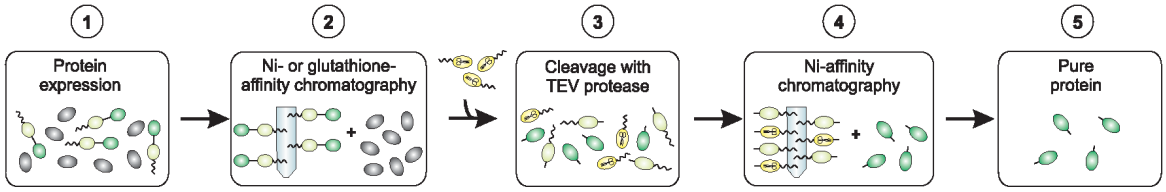

Figure 3.2: Protein purification. Dark green spheres represent the target protein, light-green speres the fusion tag, while the TEV protease is depicted in yellow and by a pair of scissores in the spheres. Dark grey spheres depict soluble, but unwanted *E. coli* proteins. 1. Soluble proteins contained in the cell lysate supernatant. 2. Separation of His-tagged target proteins from non-His-tagged E. coli proteins. 3. Cleavage of the eluted proteins to remove the His-tag from the target proteins. 4. Separation of the non-tagged target protein from the His-tagged cleavage products. 5. Solution of the pure non-tagged target protein, which can be used for NMR studies.

3.2.1 Transformation

To introduce the expression vector into the host cells, plasmid DNA is added on ice to ≈75 μl DMSO competent *E. coli* BL21(DE3) cells (Inoue *et al.*, 1990) and incubated on ice for 10 min. To enhance the transformation efficiency the DNA-cell mixture is heated for 1 min. to 42 °C and subsequently cooled on ice for another minute. Following this heat-shock the mixture is added to 1 ml LB medium and incubated at 37 °C for half an hour in the absence of antibiotics to allow the cells to recover. The cells are then spun down by short centrifugation and the cell pellet is resuspended in ≈100 μl supernatant. After plating the cells on antibiotics containing agar plates, cell colonies are grown over night at 37 °C.

3.2.2 Protein expression and purification

For a pre-culture 50 ml LB- or minimal medium were inocculated with a single colony from the agar plates and grown for 10 hours at 37°C. 10 ml dense pre-culture were then added to 800 ml LB-, ^{15}N-, or ^{13}C,^{15}N-labelled minimal medium. To overexpress the recombinant proteins the cells were induced with IPTG (0.5–1 mM final concentration) at an optical density of ≈0.8 at 600 nm and harvested 2–4 h after induction (step 1 in Fig. 3.2). For the *Sc.* Clf1 TPR constructs the cells were transferred to 15°C prior to IPTG induction and incubated over night.

Cell pellets were resuspended in chilled lysis buffer A, incubated on ice with 1 mg/ml DNase I, 1 mg/ml lysozyme, 1 M PMSF in a 1:1000 dilution and β-mercaptoethanol in a 1:10.000 fold dilution. The suspended cell pellets were homogenised by sonification and consequently centrifuged at 45.000 rpm (Ti70 rotor, Beckmann) and 4°C for 45 min. During centrifugation Ni-resin (Qiagen) was equilibrated with lysis buffer A. The supernatant was then added to the equilibrated Ni-resin and shaken in the cold. After an hour the supernatant was transferred together with the Ni-resin into columns and eluted (step 2 in Fig. 3.2). The Ni-resin with the His-tagged proteins bound was first washed with lysis buffer A, then with washing buffer 1 and 2 to remove detergent traces and proteins binding unspecifically to the Ni-beads. The His-tagged proteins were then eluted using elution buffer typically yielding already almost pure (> 95%) protein.

Subsequently either to lower the imidazole concentration in the buffer for the next Ni-chromatography step or to obtain the His-tagged proteins in NMR buffer, the buffer was exchanged using Nap5 or PD10 columns (purchased from Amersham) following the provider's instructions. For fusion proteins with TEV protease site, the N-His$_6$- or N-His$_6$-GST tag was removed incubating the recombinant fusion proteins with TEV protease at a final concentration of 10 μg/ml provided by G. Stier or the protein expression unit at EMBL Heidelberg (step 3 in Fig. 3.2) and applying another Ni-chromatography step (step 4 in Fig. 3.2).

The purity of the recombinant proteins was determined by SDS-PAGE (Laemmli, 1970) and mass spectrometry (Biochemical Instrumentation Programme, EMBL Heidelberg). The

NMR samples typically had a protein concentration of 1.0 mM and were dissolved in 20 mM phosphate buffer, 30 mM NaCl, 0.03% (w/v) NaN_3 in 90% H_2O / 10% D_2O or 100% D_2O. The pH of the NMR buffer varied from 5.8 to 7.2 depending on the pI of the dissolved protein. Samples were concentrated using Amicon Centricon filtre devices with 1 or 3 kDa cut-offs (Millipore Corp.) depending on the molecular weight of the protein.

3.2.3 Determination of protein concentration

Protein concentrations were determined in two ways: First, by measuring the absorbance of the protein solution at 595 nm after addition of Bradford reagent (Bradford, 1976). The obtained value was then compared to a standard curve, for which BSA (bovine serum albumine) in different known concentrations was used. Second, by measuring the absorbance of the protein solution at 280 nm. The protein concentration could then be obtained from Lambert-Beer's Law $A_{280} = \epsilon \cdot c \cdot d$, where A_{280} is the absorption at a wavelength of 280 nm, ϵ is the extinction coefficient of the respective protein obtained from its sequence using the Schepartz Lab Biopolymer Calculator (http://paris.chem.yale.edu/extinct.html), c is the protein concentration, and d is the length of the cuvette (1 cm).

3.3 NMR spectroscopy

All NMR experiments were carried out at EMBL Heidelberg on Bruker DRX NMR spectrometers (Rheinstetten, Germany) operating at 500 and 600 MHz proton frequency and equipped with triple-resonance (^1H, ^{13}C, ^{15}N) gradient probes. Pulsed field gradients, creating temporal linear inhomogeneities of the static magnetic field, were used for artifact suppression and coherence order selection. Detection of the complex FID was accomplished by quadrature detection in the direct dimension. In the indirect dimension, States-TPPI or pulsed field gradient coherence order selection combined with sensitivity enhancement was used (Kay *et al.*, 1992). Processing of the time-domain spectra was performed using the NMRPipe/NMRDraw software package (Delaglio *et al.*, 1995). Typically, a sine-bell window function was applied to the time domain data to improve spectral resolution. All dimensions were zero-filled at least to the next power of two to improve digital resolution, and linear prediction was applied when it improved the resolution in the ^{13}C dimension, except for the NOESY spectra. The processed spectra were analysed with the program XEASY (Bartels *et al.*, 1995), except for relaxation and residual dipolar coupling data.

3.3.1 Resonance assignment

For any spectral analysis one needs to attribute the resonance lines in an NMR spectrum to the nuclei in the molecule(s) of interest (resonance assignment). Proteins can be assigned based on their amino-acid sequence (sequential assignment). In a protein NMR spectrum, the most resolved region is normally the amide region. Therefore, the assignment procedure starts with the assignment of the protein backbone using a set of 3D heteronuclear triple-resonance experiments, in which the ^{1}H–^{15}N chemical shifts are correlated with ^{13}C resonances in the backbone ($^{13}C^{\alpha}$ or ^{13}C' (carbonyl) resonances) and/or the side-chain (yielding both, $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ resonances). For instance, in a CBCA(CO)NH experiment the magnetisation is transferred from the amide protons ($^{1}H^{N}$) to the neighbouring ^{15}N and further *via* the carbonyl carbon to the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ of the preceeding residue (Fig. 3.3, left panel). Comparing the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ resonances from residue i and residue i-1 with those containing only resonances from residue i-1, one can "walk" through the primary protein sequence in a straightforward manner (Fig. 3.3, right panel).

Side-chain assignments can be obtained from similar experiments correlating all side-chain proton or carbon resonances to the amide shifts of the preceeding residue (HC(CO)NH and C(CO)NH-TOCSY experiments). Moreover, side-chain assignments can be conveniently obtained from the HCCH-TOCSY experiment, which yields the same peaks as the ^{13}C-NOESY, but contains only intraresidual connectivities.

Almost complete ^{1}H, ^{15}N and ^{13}C resonance assignments were accomplished for all proteins used in this Thesis using the experiments listed in Table 3.1 or a subset of those. In addition for the Prp40 WW domain pair a $^{3}J^{h}$-HNCO experiment was used to identify HN–CO hydrogen bond across the β-strands (Cordier & Grzesiek, 1999; Sattler *et al.*, 1999).

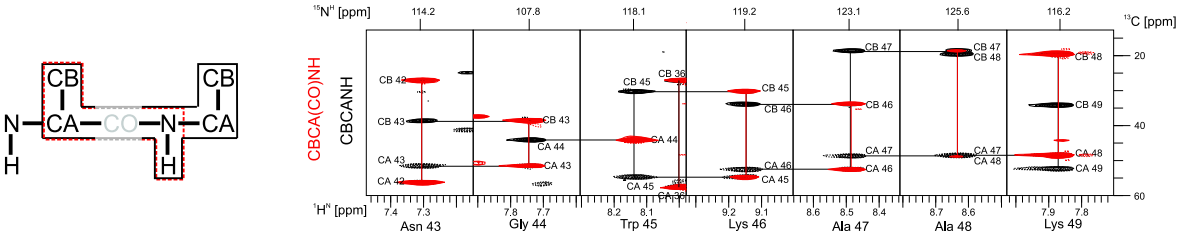

Figure 3.3: NMR experiments for the backbone resonance assignment. (a) Magnetisation transfer for the CBCA(CO)NH experiment (red) and the CBCANH (black) experiment. Since in neither case the carbonyl chemical shift is recorded, the latter is shown in grey. (b) Strips from spectra used for the sequential assignment of the Prp40 WW domain pair. Colours are as in (a).

Table 3.1: Experiments performed for the assignment of the Prp40 WW domain pair.

Experiment	Complex points acquired	Spectral width [ppm]	Assignments obtained (^{13}C, ^{15}N, ^{1}H)
HN-HSQC	256×1024	29(^{15}N), 17(^{1}H)	N(i), H^{N}(i)
HC-HSQC	200×1024	66(^{13}C), 17(^{1}H)	aliphatic C(i), H^{C}(i)
HC-HMQC	256×1024	66(^{13}C), 17(^{1}H)	C(i), H^{C}(i)
(HB)CB(CD)HD	88×1024	40(^{13}C), 17(^{1}H)	C^{β}(i), $H^{\delta,\ aro}$(i)
(HB)CB(CDCE)HE	88×1024	40(^{13}C), 17(^{1}H)	C^{β}(i), $H^{\varepsilon,\ aro}$(i)
HNCO	140×120×1024	24(^{15}N), 22(^{13}C), 17(^{1}H)	C'(i-1), N(i), H^{N}(i)
HNCA	128×100×1024	29(^{15}N), 33(^{13}C), 17(^{1}H)	C^{α}(i-1, i), N(i), H^{N}(i)
HNHA	280×136×1024	12(^{1}H), 24(^{15}N), 13(^{1}H)	H^{α}(i-1, i), N(i), H^{N}(i)
HBHA(CO)NH	130×140×1024	24(^{15}N), 17(^{1}H), 17(^{1}H)	$H^{\alpha,\beta}$(i-1), N(i), H^{N}(i)
HC(CO)NH	130×140×1024	24(^{15}N), 17(^{1}H), 17(^{1}H)	H^{ali}(i-1), N(i), H^{N}(i)
CBCANH	120×110×1024	29(^{15}N), 66(^{13}C), 17(^{1}H)	$C^{\alpha,\beta}$(i-1, i), N(i), H^{N}(i)
CBCA(CO)NH	120×124×1024	29(^{15}N), 66(^{13}C), 17(^{1}H)	$C^{\alpha,\beta}$(i-1), N(i), H^{N}(i)
C(CO)NH	170×140×1024	25(^{15}N), 70(^{13}C), 17(^{1}H)	C^{ali}(i-1), N(i), H^{N}(i)
HC(C)H	160×138×1024	8(^{1}H), 70(^{13}C), 17(^{1}H)	H^{ali}(i), C^{ali}(i), H^{ali}(i)
^{15}N-NOESY	140×238×1024	25(^{15}N), 13(^{1}H), 13(^{1}H)	H^{NOE}, N(i), H^{N}(i)
^{13}C-NOESY	256×154×1536	11(^{1}H), 66(^{13}C), 13(^{1}H)	H^{NOE}, C(i), H^{C}(i)

3.3.2 Derivation of structural restraints

For the structure calculation, geometric conformationally restrictive constraints have to be extracted from suitable NMR parameters. These include distance restraints from nuclear Overhauser effects (NOEs), hydrogen bond restraints obtained from $^{1}H/^{2}H$ exchange experiments of labile protons and scalar couplings across hydrogen bonds, dihedral angle restraints derived from three-bond scalar couplings, chemical shifts and global, distance-independent angular restraints from residual dipolar couplings, magnetic susceptibitly and rotational diffusion anisotropy.

Distance restraints

The main source of geometric information contained in the experimental NMR restraints is provided by the nuclear Overhauser effect (NOE). The nuclear Overhauser effect arises from the transfer of longitudinal magnetisation between directly coupled spins by dipole-dipole

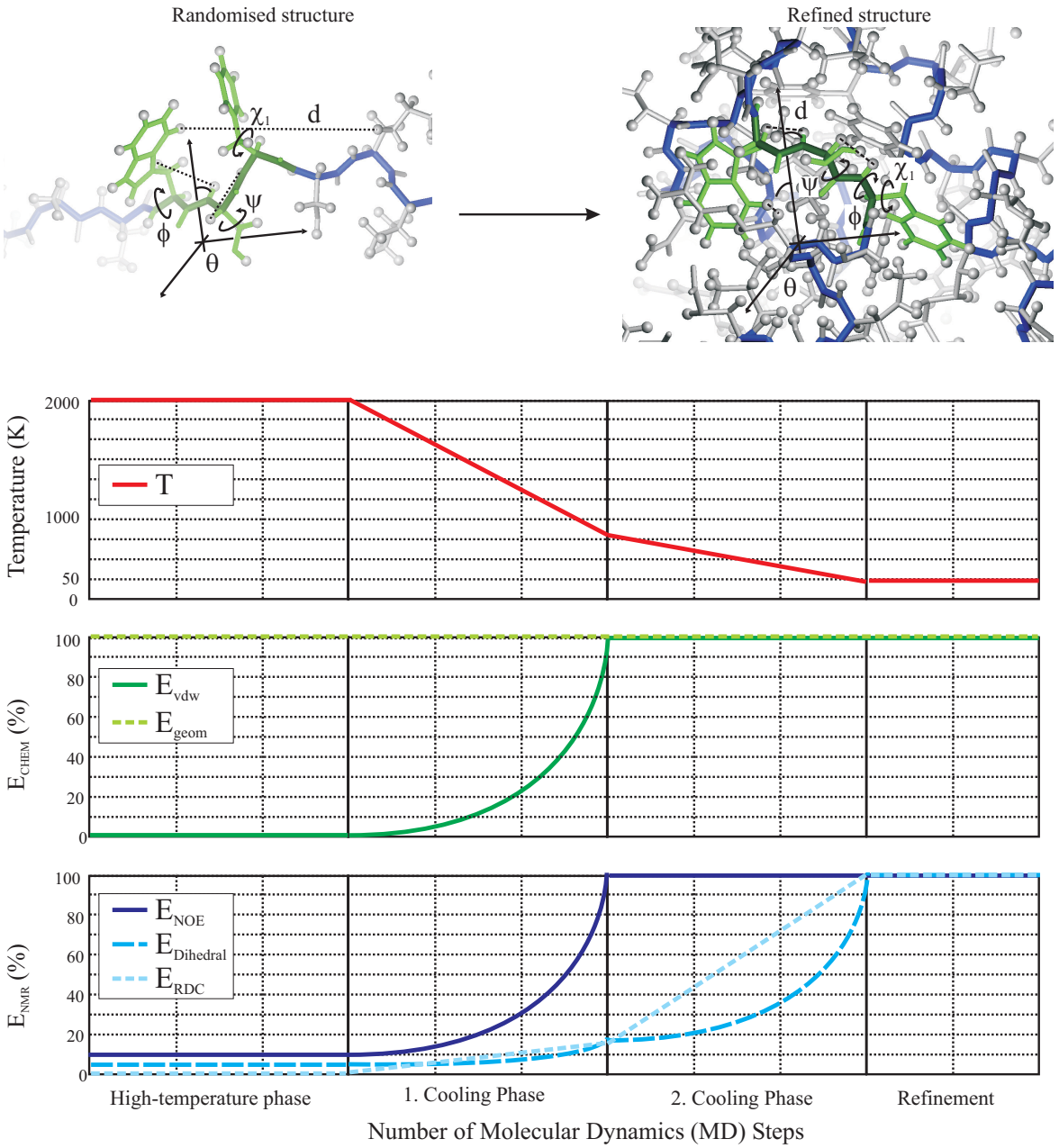

Figure 3.4: Structure refinement. Top: NMR restraints that are gradually applied during structure calculation leading from a randomised structure (left panel) to a folded structure after the refinement (right panel). d denotes inter-proton distance restraints (NOEs), ϕ, ψ and χ_1 torsion angle restraints, while θ denotes the angle of an inter-nuclear vector to an external Cartesian tensor frame as used for residual dipolar couplings. Bottom: Restrained simulated annealing applied to calculate NMR structure and the energy percentage of structural restraints during different steps of the structure calculation, which is related to the force constant of the restraint as given in (3.11) and (3.12).

cross-relaxation. The intensity of an NOE, *i.e.* the volume V of a cross-peak in a NOESY spectrum, is at short mixing times related to the distance between the involved spins as

$$V = \langle r^{-6} \rangle \cdot f(\tau_c), \tag{3.1}$$

where $\langle r^{-6} \rangle$ indicates the mean distance between the spins, which is averaged due to the flexibility of the protein. The term $f(\tau_c)$ accounts for the dependence of the magnetisation transfer on the rotational motion and the internal dynamics of the protein. By its dependence on r^{-6}, the intensity of NOEs thus falls off rapidly with increasing inter-proton distances and can therefore only be observed up to ≈5–6 Å. Despite their short-range nature NOEs are conformationally highly restrictive, particularly for protons remote in the sequence, but close in space.

The NOE intensities measured in the NOESY spectra are calibrated and used to derive ^{1}H-^{1}H distance restraints. In this Thesis, 3D ^{15}N- and ^{13}C-edited NOE experiments were used to obtain interproton distance restraints. In the case of the Prp40 WW2 domain, however, homonuclear ^{1}H-^{1}H spectra were recorded with different mixing times and at different temperatures to assign the Prp40 WW2:PSPPPVYDA complex. To derive interproton distance restraints all fully assigned peaks in the NOESY spectra were integrated using the XEASY package and calibrated using standard CNS scripts (calib.inp) used by the programme ARIA1.0 (Brünger *et al.*, 1998; Nilges & O'Donoghue, 1998). Since no stereo-specific assignment has been performed, a floating chirality approach (Folmer *et al.*, 1997) was used in the structure refinement for degenerate methylene and isopropyl groups and H^{δ} and H^{ϵ} protons of Phe and Tyr. Hydrogen bond distances were applied as indicated by $^3J(H^N,H^{\alpha})$ and ^3J(N,C') coupling constants across hydrogen bonds, $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ secondary chemical shifts and NOE patterns. The bounds for hydrogen bond restraints were 1.8–2.3 Å for H^N–O distances and 2.8–3.3 Å for N-O distances.

Residual dipolar couplings

A major development in high-resolution NMR has been the exploitation of structural information obtained from anisotropic interactions or properties of biomolecules, such as residual dipolar couplings, magnetic susceptibility and rotational diffusion anisotropy (Tolman *et al.*, 1995; Tjandra & Bax, 1997; Tjandra *et al.*, 1997; Fushman *et al.*, 1999). Since RDCs have been used extensively throughout this Thesis, the theory underlying their use in structure determination is described here in more detail.

The measurement of residual dipolar couplings (RDCs) is possible, if biomolecules are dissolved in a medium introducing a weak alignment such that the average orientation of the molecule in solution is no longer zero. Various media are now available to align biomolecules in solution, including liquid crystalline media (Tjandra & Bax, 1997), filamentous phages (Hansen *et al.*, 1998), purple membranes formed by two-dimensional crystals of bacteriorhodopsin (Koenig *et al.*, 1999; Sass *et al.*, 1999) and mechanically stressed polyacrylamide gels (Tycko *et al.*, 2000; Sass *et al.*, 2000) (Fig. 3.5 (a)). RDCs are given as the difference between the dipolar coupling constants measured in isotropic phase (J-coupling constant) and anisotropic phase (J-coupling constant + residual *direct* dipolar coupling constant). The residual dipolar coupling constant D is according to (2.27) dependent on the average angle θ between the respective internuclear vector and the direction of the static magnetic field.

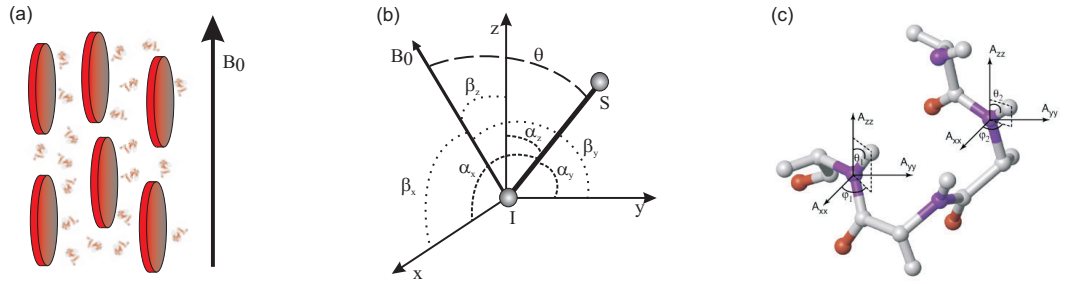

Figure 3.5: Residual dipolar couplings. (a) Introduction of orientational restrictions to the biomolecule dissolved in an alignment medium. (b) Definitions of the angles relating the molecular frame to the direction of the magnetic field used to define the order matrix. (c) Orientation of the alignment tensor with respect to the internuclear vectors.

To make use of the geometric content of RDCs in structure determination, it is according to (2.27) necessary to express the residual alignment in terms of the factor $\frac{1}{2}\langle 3\cos^2\theta - 1\rangle$. To this end, one assumes a bond vector $\vec{r}_{I,S}$ with angles α_x, α_y and α_z to the x-, y- and z-axis of the molecular frame and the angles β_x, β_y and β_z describing the orientation of the z-axis of the magnetic field with respect to the molecular frame (see Fig. 3.5 (b)), the angle θ is obtained from the dot product of $\vec{r}_{I,S}$ and $\vec{B}_0$

$$\cos\theta = \frac{\vec{B}_0\ \vec{r}_{I,S}}{|\vec{B}_0||\vec{r}_{I,S}|} = \cos\beta_x\cos\alpha_x + \cos\beta_y\cos\alpha_y + \cos\beta_z\cos\alpha_z. \tag{3.2}$$

Defining $\cos\beta_i$ and $\cos\alpha_i$ in (3.2) as b_i and a_i, respectively, yields

$$\begin{aligned}\frac{1}{2}\langle 3\cos^2\theta - 1\rangle =& \frac{3}{2}\Big[\langle b_x\rangle^2 a_x^2 + \langle b_y\rangle^2 a_y^2 + \langle b_z\rangle^2 a_z^2 \\ &+ 2\langle b_x b_y\rangle a_x a_y + 2\langle b_x b_z\rangle a_x a_z + 2\langle b_z b_y\rangle a_z a_y - \frac{1}{3}\Big].\end{aligned} \tag{3.3}$$

One can then define a so-called order matrix $\mathcal{S}$ with elements $S_{ij} = \frac{1}{2}\langle 3\cos\beta_i\cos\beta_j - \delta_{ij}\rangle$, where δ_{ij} is Kronecker's symbol ($\delta_{ij} = 1$, if $i = j$; $\delta_{ij} = 0$, if $i \neq j$). The residual alignment is then given by

$$\frac{1}{2}\langle 3\cos^2\theta - 1\rangle = \sum_{i,j=x,y,z} S_{ij}\cos\alpha_i\cos\alpha_j. \tag{3.4}$$

The order matrix $\mathcal{S}$ describes hence the average angles between the magnetic field and the bonds vectors between the interacting nuclei. Since $\mathcal{S}$ is a traceless and symmetric 3×3 matrix, the number of independent elements is five. In principle, the elements of the order matrix can therefore be defined from only five independent RDCs measured for a given molecule. Furthermore, the molecular frame can be chosen such that the order matrix is diagonal. Then the dipolar coupling constant is given by

$$D_{I,S}(\alpha_x,\alpha_y,\alpha_z) = D_{I,S}^{max}\frac{3}{2}\left[\langle b_x\rangle^2 a_x^2 + \langle b_y\rangle^2 a_y^2 + \langle b_z\rangle^2 a_z^2 - \frac{1}{3}\right], \tag{3.5}$$

where $\langle b_i \rangle^2$ corresponds to the probability of finding the i^{th} axis parallel to the magnetic field and $D_{I,S}^{max} = \frac{\mu_0}{8\pi^2} \frac{\gamma_I \gamma_S}{r^3}$. Since only the relative differences in the $\langle b_i \rangle^2$ values contribute to the RDCs, one defines the traceless, diagonal alignment tensor $\tilde{A}$ by the diagonal elements $A_{ii} = \langle b_i \rangle^2 - \frac{1}{3}$ with $|A_{zz}| > |A_{yy}| > |A_{xx}|$. Expressed by its axial and rhombic component

$$A_a = A_{zz} - \frac{A_{xx} + A_{yy}}{2} \qquad \text{and} \qquad A_r = A_{xx} - A_{yy}, \tag{3.6}$$

(3.5) can be re-written in terms of the alignment tensor

$$D_{I,S}(\theta, \phi) = D_{I,S}^{max} \cdot A_a \left[(3\cos^2 -1) + \frac{3}{2} A_r \sin^2 \theta \cos 2\phi \right], \tag{3.7}$$

where θ and ϕ are the polar angles describing the orientation of the bond vector connecting I and S with respect to the principle axis system of the alignment tensor. Consequently, to make use of the measured RDCs in the structure determination the magnitude of the axial and rhombic components of the molecular alignment tensor $\mathcal{A}$ must be determined. If the angular sampling of the internuclear vectors, for which RDCs have been measured, is sufficient A_a and A_r can be estimated from a histogram of the distribution of the RDCs values, where A_{zz} and A_{yy} are obtained by taking the average of the high and low extreme values of the RDCs and A_{xx} corresponds to the most populated value in the histogram (Clore & Gronenborn, 1998).

The geometric content of the RDCs is included in the structure calculation as an additional energy term used in the simulated annealing protocal. This equation takes the form

$$E_{dip} = k_{dip}(D_{calc} - D_{obs})^2, \tag{3.8}$$

where k_{dip} is the force constant and D_{calc} and D_{obs} are the calculated and the observed values of the residual dipolar coupling constants D, respectively. D_{calc} is computed from the angles between the bond vectors and the alignment tensor, which is represented by an external axis system comprising four atoms (O, X, Y and Z, with O at the origin) and three mutually perpendicular vectors (O-X, O-Y and O-Z) (Clore *et al.*, 1998). The orientation of this axis system is allowed to float relative to the molecule to yield the best fit between the experimental and calculated dipolar couplings. Since the RDC value is consistent with more than one orientation of the internuclear vector, RDCs are best applied at a stage in the structure calculation when already pre-folded structures based on NOEs exist.

In this Thesis, DMPC/DHPC-bicelles were used for the Prp40 WW domain pair, while a mechanically stressed polyacrylamide gel was used for the Prp40 FF1 domain and the Clf1 TPR repeat.

The bicelle stock solutions were prepared at a lipid concentration of 15% (w/v) in NMR buffer and 10% D_2O. To this end, the hygroscopic DHPC was weighed in a dry atmosphere

and dissolved in cold NMR buffer. The cold DHPC-solution was added to the solid, weighed in amount of DMPC to give a molar ratio of DHPC:DMPC ≈ 1:3 and a total lipid concentration of 150 mg/ml. This mixture was vortexed overnight at 4°C and subjected to a few cycles of warming up to 32°C and subsequent cooling at 4°C to increase the stability of the bicelles. The NMR-samples were finally prepared by diluting the bicelle stock solutions with the 0.5 mM ^{15}N-labelled sample of the non-tagged Prp40 WW domain pair to a bicelle concentration of 5% (w/v). $^{1}D_{NH}$ residual dipolar couplings were measured at 22°C (isotropic phase) and at 32°C (anisotropic phase) using a 2D spin-state-selective coherence transfer (S^{3}CT) HSQC experiment (Lerche *et al.*, 1999).

Polyacrylamide gels were casted from a stock solution containing 29.2% (w/v) acrylamide and 0.78% (w/v) N,N-methylenbiscacrylamide, which was diluted in H_2O to a final acrylamide concentration of 10% (w/v). The gel polymerisation was started by the addition of 0.1% (w/v) APS and 0.5% (w/v) TEMED in 5.0 mm (outer diameter) tubes, which were sealed with parafilm on one side. After about an hour, the polymerised gels were pushed out from the glass tubes and washed for several hours with water to remove unpolymerised acrylamide. The washed gels were cut to give final lengths of about 20 mm and subsequently soaked with the protein solution in NMR buffer in a Shigemi tube for several hours. Mechanical stress was applied by pushing the plunger of the Shigemi tube onto the gel with the plunger held in place by wrapping with parafilm. For the measurement of the RDCs a modified IPAP (in-phase, anti-phase) experiment was used tailored to suppress strong natural-abundance ^{15}N background signals from amide groups in polyacrylamide (Ishii *et al.*, 2001).

All residual dipolar coupling data were analysed using the software package CAPP/PIPP (Garrett *et al.*, 1991).

Torsion angle restraints

Torsion angles can be derived from $^3J-$coupling constants using (2.7) directly from the intensity ratios between the peaks in the decoupled and the coupled spectrum. In this Thesis, the dihedral angles ϕ and χ_1 were derived from $^3J(H^N,H^\alpha)$ and $^3J(N,C^\gamma)$ coupling constants measured by quantitative J-correlation (Kubinowa *et al.*, 1994) and spin-echo-difference experiments (Hu & Bax, 1997), respectively. The H^N,H^α coupling constant is evaluated as

$$^3J(H^N, H^\alpha) = \frac{\arctan^2(I_{cross}/I_{diag})}{\pi \cdot \delta \cdot k}, \tag{3.9}$$

where I_{cross} and I_{diag} are the intensities of the cross- and the diagonal peak, respectively, which were obtained with the programme XEASY (Bartels *et al.*, 1995), while δ is the time during which the evolution delay of coupling and k is a constant accounting for differential relaxation of the $^1H^\alpha$ and $^1H^N$ resonances. δ was set to 26 ms, while k was 1.11. For $^3J(H^N,H^\alpha)$ couplings, values larger than 8 Hz indicate $\beta-$sheets corresponding to ϕ restraints of -120° ± 40°, while values smaller than 6 Hz indicate α-helical regions corresponding to ϕ restraints of -60° ± 40°. Restraints for the backbone angles ϕ and ψ were furthermore derived from

the programme TALOS (Cornilescu *et al.*, 1999), where the experimental ϕ angles agreed well with those predicted from TALOS. However, the upper and lower restraint bounds were doubled as compared to those suggested by TALOS.

In the case of the $^3J(\mathrm{N,C}^\gamma)$ coupling, the intensity ratios between the peaks in the decoupled and the coupled spectrum are used to calculate the $^3J-$coupling constant

$$^3J(\mathrm{N,C}^\gamma) = \left[\pi/2 - \arctan^2\left[\frac{I_{coupl}/I_{ref}}{\sqrt{1-(I_{coupl}/I_{ref})^2}}\right]\right)/(\pi\cdot\delta), \tag{3.10}$$

where I_{coupl} and I_{ref} are the intensities of the coupled and decoupled spectrum, while δ is defined as in (3.9) and set to 160 ms. The side-chain angle χ_1 was restrained to 180° ± 40° for trans-rotamers ($^3J(\mathrm{N,C}^\gamma)$>1.50) and 0° ± 90° for gauche-rotamers ($^3J(\mathrm{N,C}^\gamma)$<1.00). If a dipolar coupling interaction was detected for aromatic side-chains, χ_1 was restrained to 180° ± 40°, otherwise to 0° ± 90°.

3.3.3 Structure refinement and validation

Irrespective of the algorithm used, 3D structure determination by NMR involves finding the global minimum region of a target function

$$E_{tot} = E_{NMR} + E_{chem} = E_{NOE} + E_{dih} + E_{RDC} \; + \; E_{cov(geom)} + E_{non-bonded(vdw)}. \tag{3.11}$$

E_{NMR} describes the experimental NMR restraints, *i.e.* distance restraints from NOEs, dihedral angles *e.g.* from $J-$coupling experiments and global restraints, such as residual dipolar couplings. E_{cov} includes covalent geometry restraints (bonds, angles, planarity and chirality) and $E_{non-bonded}$ van der Waals-interactions to account for atom clashes. All restraints are incorporated in the respective energy term as the sum of the weighted squares of the restraint violations and have thus the form

$$E_0 = \sum k_0(r_{obs} - r_{calc})^2, \tag{3.12}$$

where the sum runs over the number of the respective restraints, k_0 is the force constant with which the restraint is applied and r_{obs} and r_{calc} are the observed and calculated values of the restraint, respectively. The force constants used in this Thesis were 50 kcal·mol^{-1}·Å^{-2} for distance restraints, 200 kcal·mol^{-1}·rad^{-2} for torsion angle restraints and 1.4 kcal·mol^{-1}·Hz^{-2} for residual dipolar coupling restraints for the Prp40 WW domain pair and 1.0 kcal·mol^{-1}·Hz^{-2} for the Prp40 FF1 domain. The CNS E_{repel} function was used to simulate van der Waals interactions with an energy constant of 250 kcal mol^{-1}·Å^{-4} using PROLSQ van der Waals radii (Linge & Nilges, 1999) implemented in the PARALLHDG force field (Kuszewski *et al.*, 1992).

The global minimum of this total energy function comprises a set of structures (structure ensemble) that are consistent with the experimental restraints. To date, the most powerful computational method to locate the global minimum region of this target function involves the application of simulated annealing (SA) by means of molecular dynamics (MD) calculations either in Cartesian or torsion-angle space. As the potential energy surface of a protein is very complex and contains many local minima, the basic idea of simulated annealing is to overcome energy barriers by initially raising the temperature of the system and to start from different randomised extended polypeptide chains or pre-folded structures, followed by slow cooling in a step-wise manner (Fig. 3.4, top). During each temperature step the conformational space is explored using molecular dynamics sampling. While the temperature is decreased, the force constants of the restraints are increased, such that the structure ensemble slowly converges to a common fold as reflected *e.g.* by a decreased backbone rmsd (Fig. 3.4). Since bond lengths and angles are generally fixed during the calculation, these coordinates can be removed to reduce the size of the searched conformational space to allow faster sampling. Structure calculations can hence be speeded up considerably by rotating the $N/3$ torsion angles of an N atom protein in internal coordinate (torsion angle) space, instead of moving its $3N$ coordinates linearly in Cartesian space (Stein *et al.*, 1997).

As the experimental NMR data are not sufficient in their own right to determine a 3D structure of a macromolecule, they are supplemented by covalent geometry restraints (*i.e.* bond lengths, angles, planarity and chirality), which can be assumed to be known to high accuracy from small molecule work, and by a force field describing the non-bonded contacts. In its simplest representation, the non-bonded potential term can be described by a simple repulsive form to prevent atomic overlap (Lennard-Jones potential).

In this Thesis, structures were calculated using a mixed torsion and Cartesian angle dynamics simulated annealing protocol with the programs CNS 1.0 (Brünger *et al.*, 1998) and ARIA 1.0 (Nilges & O'Donoghue, 1998) optimised in terms of force constants and cooling steps for the use of RDCs (if used) to enhance the convergence of the structure ensemble. The SA protocol consisted of four stages (see Fig. 3.4): a high-temperture torsion-angle dynamics (TAD) stage to randomise the structures at 2.000 K (2.000 steps); a first Cartesian dynamics cooling phase, in which the temperature is decreased from 2.000 to 1.000 K in 5.000 steps for structure calculations without RDCs and in 9.000 steps for structure calculations using RDCs. In this phase force constant of distance restraints is increased from 10 to 50 kcal·mol^{-1}·Å^{-2}, for dihedral restraints from 5 to 25 kcal·mol^{-1}·rad^{-2} and for residual dipolar couplings from 0 to 0.2 kcal·mol^{-1}·Hz^{-2}; a second Cartesian dynamics cooling phase, in which the temperature is decreased from 1.000 to 50 K in 2.000 steps for structure calculations without RDCs and in 4.000 steps for structure calculations using RDCs. In this phase force constant of dihedral restraints from 25 to 200 kcal·mol^{-1}·rad^{-2} and for residual dipolar couplings from 0.2 to 1.4 kcal·mol^{-1}·Hz^{-2} for the Prp40 WW domain pair and to 1.0 kcal·mol^{-1}·Hz^{-2} for the Prp40 FF1 domain; and a refinement phase at 50 K consisting

of 2.000 steps. During each of the eight iterations 20 structures were calculated with ARIA using the 10 lowest energy ("best") structures of each iteration as templates for the following iteration. The seven best structures were used for violation analysis. No ambiguous NOEs were used in the structure calculations. The quality of the final structures was assessed using the program PROCHECK-NMR (Laskowski *et al.*, 1996).

3.3.4 Chemical shift mapping

As the chemical shift of a nucleus is very sensitive to its electronic surrounding, chemical shift changes upon environmental changes, such as ligand binding, can be monitored and thereby yield directly the nuclei involved in the interaction. In biomolecular NMR chemical shift changes are widely used to map binding sites. In this Thesis, this method has been used to characterise the ligand recognition of the Prp40 WW domain pair and the interaction between the Prp40 FF1 domain with the N-terminal TPR motif of the spliceosomal scaffolding protein Clf1.

Samples for NMR titrations contained typically 0.5 mM ^{15}N-labelled protein in the same buffer as used for the structure determination. Chemical shift perturbation experiments were performed at 12°C using 2D ^{15}H,^{15}N-HSQC experiments. Unlabelled peptide solutions were prepared in typically 50-fold concentration as compared to the ^{15}N-labelled Prp40 WW domains and added in a step-wise manner.

3.3.5 Nuclear spin relaxation experiments

Non-equilibrium magnetisation is driven back to thermal equilibrium by various relaxation mechanisms. Nuclear spin relaxation is described by two time constants: T_1 for the return of longitudinal magnetisation (z-magnetisation) and T_2 for the decay of transverse magnetisation (x, y-magnetisation). As the molecule rotates in solution, the orientation of the local magnetic field at the site of the nucleus varies randomly with respect to the static magnetic field and the "memory" for its initial orientation is lost with time. The latter is described by the so-called auto-correlation function, which decays exponentially with the rotational correlation time τ_c of the molecule. Fourier transformation of the auto-correlation function into frequency domain yields the spectral density function

$$J(\omega) = \frac{2\tau_c}{5(1+\omega^2\tau_c^2)}, \tag{3.13}$$

which can be directly related to experimental spin relaxation rates.

The most frequently measured relaxation rates in biomolecular NMR are the ^{15}Nlongitudinal and transverse relaxation rates and the steady-state $\{^1$H$\}$-^{15}N heteronuclear NOE, which can

be expressed as linear combinations of the spectral density functions of the $^{15}N^{H}$ and ^{1}H nuclei

$$R_1 = \frac{1}{T_1} = \frac{d^2}{4}\left[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)\right] + c^2 J(\omega_N) \tag{3.14}$$

$$R_2 = \frac{1}{T_2} = \frac{d^2}{8}\left[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)\right] + \frac{c^2}{6}\left[3J(\omega_N) + 4J(0)\right] + R_{ex} \tag{3.15}$$

$$\eta_{\mathrm{HN}} = \frac{d^2}{4}\left[6J(\omega_H) + 6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)\right] \tag{3.16}$$

where $d = (\mu_0/8\pi^2)h\gamma_H\gamma_N\langle r_{\mathrm{HN}}\rangle^{-3}$, $c = \gamma_N B_0 \Delta\sigma/3$, μ_0 is the permittivity of free space, $\Delta\sigma$ is axially symmetric chemical shift anisotropy (CSA) and R_{ex} is the chemical or conformational exchange contribution to R_2.

At a given frequency reorientational motions match the Larmor frequency and induce most effectively relaxation. Since the rotational correlation times of small to medium-sized biomolecules fall in the lower ns time-scale at ambient temperature, the dispersion of $J(\omega)$ can be used to obtain information about the dynamics of biomolecules. Moreover, in the case of anisotropic tumbling the local rotational correlation times can be related to an anisotropic diffusion tensor yielding long-range restraints for structure calculation and validation.

Rotational diffusion tensor

An overall rotational diffusion tensor can be estimated from the ratio of the transverse and longitudinal relaxation rates (R_2/R_1). This ratio is directly related to the local rotational correlation time $\tau_{c,i}$ for nuclei, which are not subject to fast internal motions or conformational exchange. The overall diffusion tensor is a second rank tensor, which can be represented by a 3×3 matrix and is defined by six parameters, such as the lengths of the three principal axes D_{xx}, D_{yy} and D_{zz} and the Euler angles α, β and γ determining the orientation of the diffusion frame with respect to the molecular coordinate frame. In the case of anisotropic rotational diffusion, $\tau_{c,i}$ depends on the directions of the internuclear vectors (*e.g.* ^{1}H–^{15}N-vector) relative to the diffusion tensor and hence contains *a priori* long-range information about the orientation of bond vectors, which can be used for structure calculation and validation. This strategy has been used in Chapter 5 to assess the three-dimensional structure and relative domain orientation of the Prp40 WW domains. If the molecule, however, tumbles isotropically all orientational information is lost, since $\tau_{c,i}$ assumes the same value for all internuclear vectors.

^{15}N relaxation measurements

Heteronuclear $\{^{1}H\}$-^{15}N NOEs as well as longitudinal (R_1) and transverse (R_2) ^{15}N relaxation rates were measured using standard two-dimensional methods (Farrow *et al.*, 1994). For the Prp40 WW domain pair, ^{15}N relaxation data were recorded at 600.13 MHz proton frequency

and 22°C using a 1.0 mM ^{15}N-labelled sample of the non-tagged protein. The relaxation delays were applied in an interleaved fashion and set to 20 ms (run twice), 120 ms, 240 ms, 350 ms, 500 ms, 710 ms, 1.04 s, 1.4 s, 1.8 s for R_1 measurements and [1 (run twice), 2, 3, 4, 6, 8, 10, 12, 14]×15.8176 ms for R_2 measurements.

For any of the proteins used, the heteronuclear NOE experiments were run twice in an interleaved fashion with and without (reference experiment) proton saturation during the recovery delay. Errors in the peak intensities were estimated from the average base-line noise in the spectra of repeated experiments. The relaxation experiments were analysed using the program NMRView5.0.3 (Johnson & Blevins, 1994). Relaxation rates were obtained from non-linear least-squares fits using in-house developed software. The peak intensities were fitted to a two-parameter decaying exponential function of the form $I = I_0 \cdot e^{-R\tau}$, where I is the peak intensity at a given relaxation delay, τ is the relaxation delay itself, while the fit parameters I_0 and R are the peak intensity at time 0 and the relaxation rate, respectively. Errors in the fit were estimated using Monte Carlo methods with a 65% confidence interval. $\{^1$H$\}$-^{15}N NOEs were determined as the peak intensity ratio between the reference and the saturation experiment using NMRView5.0.3 (Johnson & Blevins, 1994) and are the average of two measurements. Residues subject to fast internal motions or chemical exchange were identified as described (Pawley *et al.*, 2001).

The respective diffusion tensor was determined by fitting the experimental ^{15}N R_2/R_1 ratios to the seven lowest energy structures using the program TENSOR v2.0 (Dosset *et al.*, 2000).

3.4 Other methods

Sequence alignments were performed using the program Clustal_X (Thompson *et al.*, 1997). Homology models were prepared using MODELLER (Sali & Blundell, 1993) and energy-minimised with the CHARMM (Brooks *et al.*, 1983) function implemented in the QUANTA programme package (Molecular Simulations Inc., 1997).

Figures of protein structures were prepared with the programme MOLMOL (Koradi *et al.*, 1996), which was also used to superimpose structure ensembles and homology models and to calculate protein backbone and heavy atom root mean square deviations. Protein surface representations were generated using GRASP (Nicholls *et al.*, 1993), while "mutations" in peptide coordinates were introduced using the "mutate" function in the Swiss-PdbViewer programme (Guex & Peitsch, 1997).

3.5 References

Bartels, C., Xia, T.-H., Billeter, M., Güntert, P. & Wüthrich, K. (1995). The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR,* **5**, 1–10.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–54.

Brooks, B. R., Bruccoleri, R. E., Olafsson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* **4**, 187–217.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D,* **54**, 905–921.

Clore, G. M. & Gronenborn, A. M. (1998). A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. *J. Magn. Res.* **133**, 216–221.

Clore, G. M., Gronenborn, A. M. & Tjandra, N. (1998). Direct refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J. Magn. Res.* **131**, 159–162.

Cordier, F. & Grzesiek, S. (1999). Direct observation of hydrogen bonds in proteins by interresidue $^{3h}J_{N,C'}$ scalar couplings. *J. Am. Chem. Soc.* **121**, 1601–1602.

Cornilescu, G., Delaglio, F. & Bax, A. (1999). Protein backbone angle restraints from searching a database for protein chemical shifts and sequence homology. *J. Biomol. NMR,* **13**, 289–302.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR,* **6**, 277–293.

Dosset, P., Hus, J. C., Blackledge, M. & Marion, D. (2000). Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR,* **16**, 23–28.

Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D. & Kay, L. E. (1994). Backbone dynamics of a free and a phosphopeptide-complexed Src Homology 2 domain studied by 15N NMR relaxation. *Biochemistry,* **33**, 5984–6003.

Folmer, R. H., Hilbers, C. W., Konings, R. N. & Nilges, M. (1997). Floating stereospecific assignment revisited: application to an 18 kDa protein and comparison with J-coupling data. *J. Biomol. NMR,* **9**, 245–258.

Fushman, D., Xu, R. & Cowburn, D. (1999). Direct determination of interdomain orientation on ligation: Use of the orientational dependence of 15N NMR relaxation in Abl SH(32). *Biochemistry,* **38**, 10225–10250.

Garrett, D. S., Powers, R., Gronenborn, A. M. & Clore, G. M. (1991). A common sense approach to peak picking two-, three- and four-dimensional spectra using automatic computer analysis of contour diagrams. *J. Magn. Res.* **95**, 214–20.

Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis,* **18** (15), 2714–23.

Hansen, M. R., Mueller, L. & Pardi, A. (1998). Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.* **5**, 1065–74.

Hu, J. & Bax, A. (1997). χ_1 angle information from a simple two-dimensional NMR experiment that identifies trans $^3J_{N,C_\gamma}$ couplings in isotopically enriched proteins. *J. Biomol. NMR,* **9**, 323–328.

Inoue, H., Nojima, H. & Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. *Gene,* **96**, 23–8.

Ishii, Y., Markus, M. A. & Tycko, R. (2001). Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel. *J. Biomol. NMR,* **21**, 141–51.

Johnson, B. A. & Blevins, R. A. (1994). NMRView: a computer-program for the visualization and analysis of NMR data. *J. Biomol. NMR,* **4**, 603–614.

Kay, L. E., Keifer, P. & Saarinen, T. (1992). Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663–5.

Koenig, B. W., Hu, J. S., Ottiger, M., Bose, S., Hendler, R. W. & A., B. (1999). NMR Measurement of dipolar couplings in proteins aligned by transient binding to purple membrane fragments. *J. Am. Chem. Soc.* **121**, 1385–6.

Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis macromolecular structures. *J. Mol. Graphics,* **14**, 51–55.

Kubinowa, H., Grzesiek, S., Delaglio, F. & Bax, A. (1994). Measurement of HN-Ha J-couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods. *J. Biomol. NMR,* **4**, 871–878.

Kuszewski, J., Nilges, M. & Brünger, A. (1992). Sampling an efficiency of metrix matrix distance geometry: a novel partial metrization algorithm. *J. Biol. NMR,* **2**, 33–56.

Laemmli, U. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* **227**, 680–9.

Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR,* **8** (4), 477–86.

Lerche, M. H., Meissner, A., Poulsen, F. M. & Sørensen, O. W. (1999). Pulse sequences for measurement of one-bond (15)N-(1)H coupling constants in the protein backbone. *J. Magn. Res.* **140**, 259–263.

Linge, J. P. & Nilges, M. (1999). Influence of non-bonded parameters on the quality of NMR structures: a new force field for NMR structure calculation. *J. Biomol. NMR,* **13**, 51–59.

Nicholls, A., Bharadwaj, R. & Honig, B. (1993). GRASP: Graphical representation and analysis of surface properties. *Biophys. J.* **64**, 166–170.

Nilges, M. & O'Donoghue, S. I. (1998). Ambiguous NOEs and automated NOE assignment. *Progr. NMR Spectr.* **32**, 107–139.

Pawley, N. H., Wang, C., Koide, S. & Nicholson, L. K. (2001). An improved method for distinguishing between anisotropic tumbling and chemical exchange in analysis of 15N relaxation parameters. *J. Biomol. NMR,* **20** (2), 149–165.

Sali, A. & Blundell, T. (1993). Comparative protein modelling by satisfaction of spacial restraints. *J. Mol. Biol.* **234**, 779–815.

Sass, H., Cordier, F., Hoffmann, A., Rogowski, M., Cousin, A., Omichinski, J., Löwen, H. & Grzesiek, S. (1999). Purple membrane induced alignment of biological macromolecules in the magnetic field. *J. Am. Chem. Soc.* **121**, 2047–2055.

Sass, H. J., Musco, G., Stahl, S. J., Wingfield, P. T. & Grzesiek, S. (2000). Solution NMR of proteins within polyacrylamide gels: diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. *J. Biomol. NMR,* **18**, 303–9.

Sattler, M., Schleucher, J. & Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. NMR Spectrosc.* **34**, 93–158.

Stein, E. G., Rice, L. M. & Brünger, A. T. (1997). Torsion-angle dynamics as a new efficient tool for NMR structure calculation. *J. Magn. Res.* **124**, 154–64.

Studier, F. W., Rosenberg, A. H., Dun, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Meth. Enzymol.* **185**, 60–89.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25** (24), 4876–82.

Tjandra, N. & Bax, A. (1997). Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science,* **278**, 1111–1114.

Tjandra, N., Garrett, D. S., Gronenborn, A. M., Bax, A. & Clore, G. M. (1997). Defining long range order in NMR structure determination from the dependence of heteronuclear relaxation times on rotational diffusion anisotropy. *Nat. Struct. Biol.* **4**, 443–9.

Tolman, J. R., Flanagan, J. M., Kennedy, M. A. & Prestegard, J. H. (1995). Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. *Proc. Natl. Acad. Sci. USA,* **92**, 9279.

Tycko, R., Blanco, F. & Ishii, Y. (2000). Alignment of biopolymers in strained gels: A new way to create detectable dipole–dipole couplings in high-resolution biomolecular NMR. *J. Am. Chem. Soc.* **122**, 9340–9341.