αB-crystallin: Insights into a small heat shock protein
with NMR-spectroscopy

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry and Pharmacy

of Freie Universität Berlin

by

Lea Johanna Münkemer, geb. Lauer

Diploma-Chemist

March 2019

This work was conducted from April 2013 to March 2019 at the Leibniz-Forschungsinstitut für molekulare Pharmakologie (FMP) under the supervision of Dr. Barth-Jan van Rossum and Prof. Dr. Hartmut Oschkinat.

Reviewers: 1. Prof. Dr. Hartmut Oschkinat, FU Berlin
              2. Prof. Dr. Bernd Reif, TU München

Date of Disputation: 11. Juni 2019

"It is clear that neither model offers a satisfactory description of the aggregate. This raises the question of whether it will be possible, at all, to explain the many apparently conflicting observations in the literature in terms of a unique structure or set of structures in which the position of any subunit can be described with a closed set of simple rules for translation and rotation."

R. Augusteyn, J. Koretz, 1987

Abbreviations

ppm .... *parts per million*

PRE .... *paramagnetic relaxation enhancement*

rf .... *radio frequency*

SAXS .... *small angle X-ray scattering*

sHSP .... *small heat shock protein*

SSB .... *sine bell shift*

TOTAPOL .... *1-(TEMPO-4-oxy)-3-(TEMPO-4-amino)propan-2-ol*

Table of contents

1 Introduction

1.1 Introduction to αB-crystallin

Proteins are the key players in every aspect of living organisms and evolution generated a huge machinery for homeostasis[1,2], which describes all processes which are involved in keeping the balance between proper protein folding, unfolding and aggregation. The protein class of the small heat shock proteins (sHSP) is ubiquitous and members acts as passive, energy-independent chaperones[3–8]. These proteins are responsible for keeping un- or misfolded proteins in solution and thereby preventing aggregation[5,8–13].

The 20 kDa protein αB-crystallin (αB) belongs to this class of sHSPs. It was initially discovered in the eye-lens[14] as a subunit of α-crystallin, an oligomeric ensemble comprised of the two homologues αB- and αA-crystallin (αA) subunits in a ratio of 1:3[15,16]. Together with the β- and γ-crystallins they make up to 90% of the eye-lens protein of vertebrates[17,18]. Cataract, a disease developed with ongoing lens-aging[18], is the major reason for acquired blindness worldwide[19] and found to be related to advanced post-translational modifications of α-crystallin[20].

Though αB occurs always as mixed oligomer with the homologue αA-crystallin in the eye lens, unlike the homologue, αB is found in every other tissue of the human body[21]. In neurodegenerative diseases like Parkinson[22,23], Lewy body-disease[24,25], Alzheimer's disease[26,27], Alexander's disease[28–32] or certain cancer types[33–35] as well as in states of cellular stress[31] αB is found to be highly upregulated. Furthermore, the chaperone is associated with cardio myopathies[29–31], hypoxia[36], multiple sclerosis[37,38] and the expression may be altered for patients suffering from tuberous sclerosis[39]. In addition, the occurrence of αB-crystallin is related to the ubiquitine proteasome pathway and apoptosis of cells[14,40].

Mutations of the protein are often related to diseases or the early onset of them compared to the age-related disease-state. E.g., the R120G-mutant is linked to cardiomyopathies and the early onset of cataract[41,42]. Hereditary cataract is moreover found to be caused by the mutation αB D140N[43].

1.1.1 Architecture of αB-crystallin and oligomeric distribution

The sHSP αB-crystallin consists of 175 AA, 20 kDa respectively. It can be divided into three domains (**Figure 1**): First, a hydrophobic N-terminal domain (NTD) of approximately 60 residues, which contains the three major phosphorylation sites (Ser19, Ser45 and Ser59) as discussed in 1.1.3. Second, a conserved domain of about 90 residues, the α-crystallin domain (ACD)[3,44,45] and third, a flexible C-terminal domain (CTD) of approximately 25 residues. The short C-terminus includes a highly conserved IXI-motif in position 159-161[6,46], followed by a very hydrophilic extension (C-terminal extension, CTE) that provides the protein its excellent solubility properties. In fact, αB-crystallin is one of the most soluble proteins known, an indispensable feature in the eye-lens where the concentration of α-crystallin is up to 450 mg/mL[47].

| NTD | α-crystallin domain ACD | CTD |
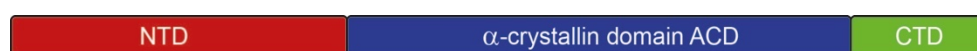
Figure 1: General composition of αB-crystallin; red: N-terminal domain (NTD), blue: conserved α-crystallin domain (ACD) and green: C-terminal domain (CTD). Sizes of domains are depicted in relation to the full length of the protein (175 residues).

The protein occurs in solution as a distribution of oligomers, which cannot be isolated from each other. The sizes vary from 10 to 40-mers with a maximum around 24-28 monomers per oligomer depending on the type of measurement and concentration[48–50] (**Figure 2**). In general, particles with an even number of subunits are favored over those with an odd number[51]. Due to the variable sizes and the heterogeneous structure, αB-crystallin is generally described as polydisperse. Short constructs of the protein may occur in ensembles of discrete sizes, like dimers, tetramers or hexamers (see 1.1.1.2).
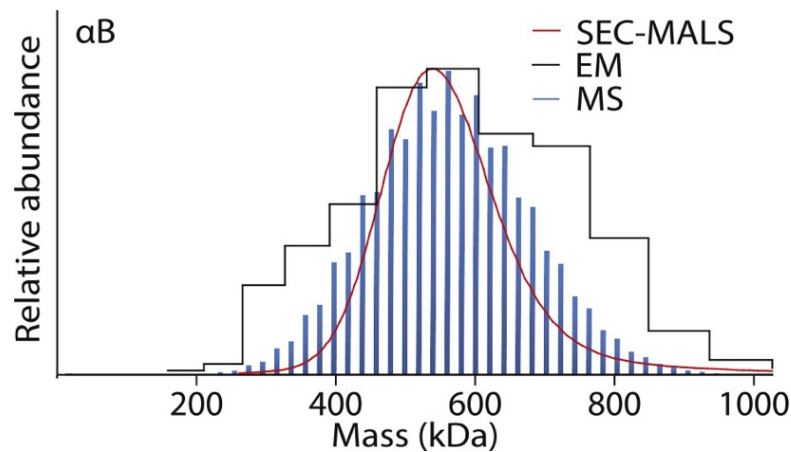
Figure 2: Distribution of oligomeric masses of αB-crystallin assessed with different methods (SEC-MALS, EM and MS) according to Hochberg *et al.*[48].

1.1.1.1 Early structure-models of α-crystallin

Given the fact that αB-crystallin was discovered as a mixed oligomer with αA-crystallin, the first attempts of describing a structure aimed for the hetero-oligomer α-crystallin. Biochemical data revealed the ratio of 1:3 of the two subunits and first results showed species of growing molecular weight upon aging and higher glucose concentration in bovine lens α-crystallin[52]. Siezen *et al.* made an attempt to describe the ensemble as a three layer model based on their investigations of temperature and pH on the particle size[53] (**Figure 3 A**). They examined velocity effects and sedimentation coefficients in presence and absence of calcium-ions and performed electron-microscopy (EM) on the hetero-oligomer.

In 1986, Tardieu *et al.* acknowledged the need of accounting the variety of sizes and molecular weight species observed under the various physicochemical conditions, by suggesting a model built of multiple layers. The core-layer contains either the αA- or αB-subunit, surrounded by a definite number of protomers of the other subunit in the second layer, which, in turn, is enclosed by a third layer of the monomers present in the core. This resulted in a symmetrical model of a quaternary structure comprising both α-crystallin subunits in the second and third layer with a limited number of subunit-subunit contacts[54] (**Figure 3 B**).

Augusteyn and Koretz developed a model based on the biochemistry of the subunits and the 'sharp borders of hydrophobicity and hydrophilicity in the sequence'[55]. They suggested that α-crystallin would be the first micelle-structured protein known, since this type of arrangement of subunits would be in agreement with all experimental

data collected until that point of time, including dependence of particle size in vitro on temperature, ionic strength, specific ion effects and increased concentration leading to increased particle sizes.

The multilayered type of substructure and the idea of a micelle like arrangement of particles was taken up also by Walsh *et al.*, who conducted differential scanning calorimetry on α-crystallin and found two thermal events which were believed to derive from subunit assembly[56]. Their model (**Figure 3 D**) comprised an inner micelle layer with cub-octahedral symmetry, a second similar layer with 8 subunits facing with their hydrophobic parts to the core and thus giving a particle with octahedral symmetry and a third outer layer, which adds up to 42 subunits in total.

In 1993, Wistow stated it to be neither unreasonable to suppose a single structure nor possible to reconcile all the conflicting biophysical and biochemical data[57]. His models (**Figure 3 C**), a cube and a dodecahedron, were developed with the assumption that the internal protein sequence symmetry and gene structure suggest a two-domain construction. According to Wistows' division, the N-terminal domains contain the hydrophobic parts of the sequences in the beginning and the C-terminal domains include the conserved crystallin-domain and the hydrophilic tails of the proteins. For the conserved, hydrophilic part four contact areas are postulated, marked in **Figure 3 C** with a, b, c and d. The N-terminal parts are believed to multimerize in less defined ways. The suggested cube is to be built out of 24 subunits, whereas the rhombic dodecahedron would contain 48 monomers, which would fit better to the experimental mass of the hetero-oligomer.

Shortly after that in the early 90s, the chaperoning function of the α-crystallin was discovered[44] and Carver was able to show the highly flexible character of the C-termini of both subunits[58,59], αA-crystallin and αB-crystallin, by performing TOCSY and ROESY type solution nuclear magnetic resonance (NMR) measurements. Experiments on the homo- and the hetero-oligomers resulted in the assignment of the last 10 residues for αB and 8 residues for αA and it was shown that the C-termini do not interact with each other in the hetero-ensemble. Two isoleucine-signals could be identified as well as an alanine but these could not be assigned sequentially, therefore an attribution to the N-terminal residues 3 – IAI – 5 was hypothesized. Furthermore, urea-based denaturation experiments monitored by NMR affirmed a

two-domain structure as unfolding occurs in two steps[58]. As a result, Carver *et al.* suggested a 'chaperone' structure for the system[60] (**Figure 3 E**). They described a funnel built out of four rings, each comprising only one type of protomers, in the order B-A-A-B. The subunits were depicted as balls with unknown structures, and flexible and solvent accessible C-termini cover the cavity. Hydrophobic parts are believed to be on the inside of the funnel and the phosphorylation sites in the N-terminus surface-exposed in order to be assessable for post-translational modifications. Particles with average masses other than 800 kDa could derive from detached rings or constructs with less than 20 monomers per ring.
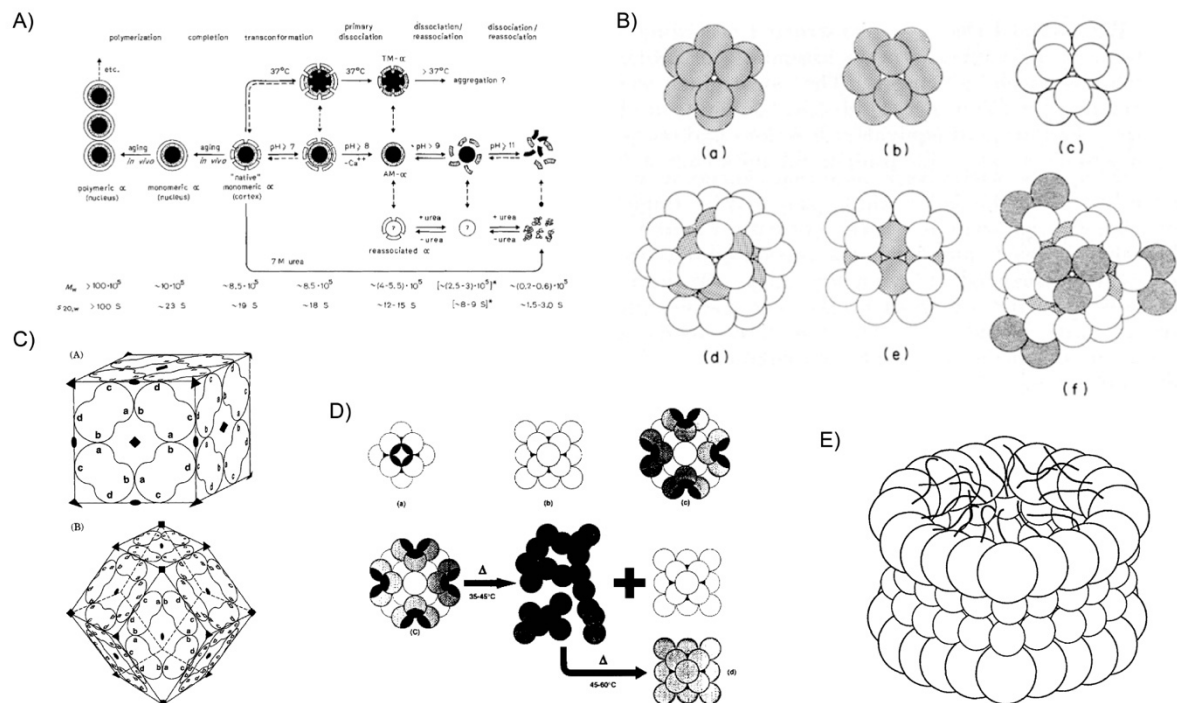
Figure 3: Early models for oligomers of α-crystallin: A) Three-layer model from Siezen and Bindels proposed in 1980[53],that was developed based on velocity effects and sedimentation coefficients. **B)** Symmetrical quaternary structure with limitation of subunit-subunit contacts from Tardieu *et al.* proposed in 1986[54]. **C)** Models based on two-domain interactions of NTD and CTD, resulting in a 24mer as cube and a 48mer as rhombic dodecahedron, models by Wistow published in 1993[57]. **D)** Layer-model by Walsh *et al.* suggested in 1991, with oriented subunits according to the degree of hydrophobicity of the domains[56]. **E)** Chaperone-structure by Carver *et al.* from 1994, with flexible C-termini covering the funnel[60].

1.1.1.2 Structure-models of αB-crystallin

As the hetero-oligomer was present only in the eye-lens, but the αB subunit was found in many other tissues, the further research essentially concentrated on the more ubiquitous subunit of α-crystallin, αB-crystallin. Incited by method development in both electron microscopy and NMR-spectroscopy, leading to improved resolution and the possibility of recombinant protein production, the number of publications mentioning αB-crystallin increased greatly after the suggestions of the theoretical models for α-crystallin. Since till date the full-length protein was not amenable for crystallization, a combination of methods including EM, NMR and X-ray crystallography on short constructs for characterization of the polydisperse and oligomeric ensemble is indispensable.

Early electron microscopy studies showed that αB-crystallin has a quaternary structure with roughly spherical shape and a central cavity[61]. This picture of a hollow spherical particle was affirmed later by cryo-EM at 34 Å resolution[62] (**Figure 4 A**).

Muchowski *et al*. presented a homology-model[63] based on the 2.9 Å resolution crystal structure of a homologue derived from *Methanocaldococcus jannaschii* (formerly *Methanococcus jannaschii*), a hyperthermophilic archaeon[64] (**Figure 4 B**). The small heat shock protein MjHSP16.5 was the first protein of this class for which a structure could be elucidated. It comprises an α-core domain, similar to αB-crystallin with a sequence identity between the two proteins of 19%, whereas the chemical similarity in residues of the core is around 51%. Yet, due to the lacking homology of the terminal regions, the model for αB-crystallin comprises mainly a structured core domain with a β-sandwich fold, flanked by unstructured termini. Furthermore, it does not display contacts of residues 114-118 of the core domain which were observed via EPR-measurements[65]. Feil *et al*. investigated the α-core domain of αB-crystallin, using the short construct comprising residues 57-157, and which forms only dimers. By synchrotron radiation X-ray scattering and rigid body refinement, they derived a structure for the α-crystallin core domain differing from the homology model and the MjHSP16.5 structure[66]. No reasonable fit of the experimental X-ray solution-scattering data to the structure of the homologue could be obtained, therefore rigid body refinement was applied. This led to a model which was in line

with the spin-labelling results from Koteiche[65] and could be fitted to the scattering curves. This model is the first to comprise a dimeric organization in which an antiparallel β-sheet arrangement of two subunits form the interface within the core-domain dimer by modelling the region from residue 105 to 116 as a strand-turn-strand motif named β6.1 and β6.2 (**Figure 4 C**).
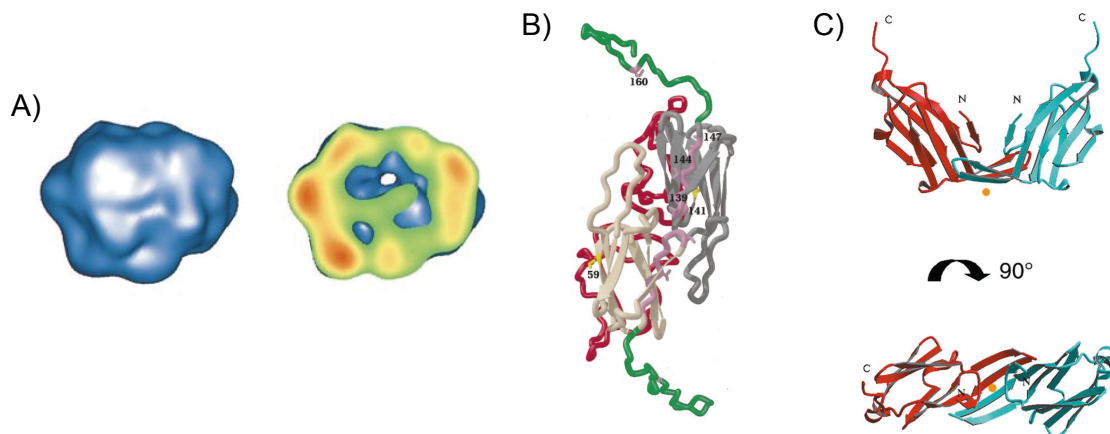
Figure 4: A) Cryo-EM structure of αB-crystallin with 34 Å resolution published in 2000 by Haley *et al.*[62] depicting the hollow spherical particle shape. B) Homology model of the αB-crystallin core domain by Muchowski *et al.*[63] (1999) based on homology with MjHSP16.5 from *Methanocaldococcus jannaschii*, a hyperthermophilic archaeon. C) Model by Feil *et al.*[66] (2001) of the dimeric short construct comprising the α-core domain based on synchrotron radiation X-ray scattering and rigid body refinement. The model includes the antiparallel arrangement of β-sheets in the interface of the two protomers.

The structured nature of the α-core domain also became clear from magic-angel spinning (MAS) solid-state NMR results of αB-crystallin[67], where chemical shift resonances were assigned for residues 64-162 in measurements on the full-length protein. Using the fact that chemical shift resonances are very sensitive monitors for the secondary structure of a protein, analysis of the NMR-data revealed six β-strands for the α-crystallin domain, with residues 113 to 123 forming an extended β-strand, named β6+7 (**Figure 5**). The numbering of the six β-strands was carried out in line with the convention according to the alignment to known structures of the sHSPs wHSP16.9, Tsp36 and MjHSP16.5. The assignment could be affirmed by solution NMR-measurements for the isolated core-part of the protein, which forms dimers in solution and results in comparable chemical shifts as the solid-state measurements.

7

As conclusion of the chemical shift similarities between the different NMR-methods the dimer interface (DI) was determined. It is formed by the antiparallel arrangement of the long β6+7-strands from two subunits, which is in accordance with the earlier EPR-results and the model from rigid-body refinement indicating such a dimer-arrangement[65,66]. Following subsequent MAS solid-state NMR investigations, structure calculations[68] lead to the first NMR-structure at atomic resolution for the α-crystallin domain, namely a dimer with a curvature of the extended β6+7-strands at an angle of 121°.
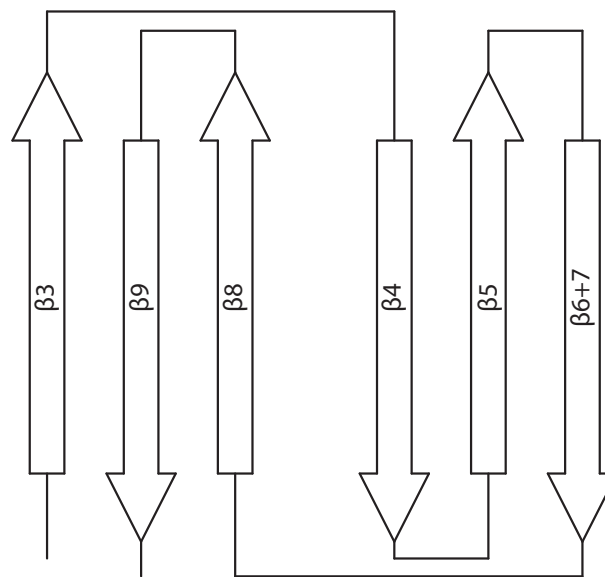
Figure 5: Schematic arrangement of β-strands in the α-crystallin domain of αB-crystallin. Figure adapted from Jehle _et al._[67].

The dimeric structure of the core domain, with the interface formed by β6+7-strands, was confirmed by X-ray crystallography on the short constructs αB-crystallin 67-157[69] and αB-crystallin 68-162[70]. Results from Laganowsky and Hochberg _et al._[70,71] showed three possible register shifts (AP$_I$, AP$_{II}$ and AP$_{III}$) for the dimer interface formed by the long antiparallel arranged β-strands. The DI is described as being pH-sensitive[67,72] and the core domain was found to lose its propensity for dimers at pH 5[73]. Further investigations revealed that the short construct αB-crystallin 68-153 exists in solution mostly in the AP$_{II}$ state, meaning the residues E117 from two

protomers face each other[71]. The AP$_{II}$ register as main conformation for the DI was confirmed with solution and solid-state NMR-spectroscopy[72,74].

Furthermore, Peschek *et al.* carried out electron microscopy measurements with recombinant αB-crystallin, in addition to biophysical characterization with methods such as dynamic light scattering (DLS) and circular dichroism (CD) spectroscopy, and presented a 3D surface model for the 24mer (**Figure 6 A**) displaying the protein as spherical particles with a large inner cavity and tetrahedral symmetry[75].

In the X-ray structure of MjHSP16.5, a C-terminal IXI-motif was found to be bound to a hydrophobic groove formed by an adjacent core domain[64]. The same situation is given for αB-crystallin as observed by X-ray crystallography on short constructs[69,70] and in the solid-state NMR-data of the full-length protein[68]. The apparent oligomer-sizes measured with DLS and small angle X-ray scattering (SAXS) increase with lower pH-values, which is interpreted as loosening of the C-termini of the protein and therefore resulting in bigger particles[68]. Laganowsky *et al.* showed in their work on short constructs of αB-crystallin (ABC68-162) that multimerization of the core domain is strongly concentration dependent[70]. At low concentrations of 0.2 mg/mL, the αB short construct exists mainly as monomer, at 0.4 mg/mL as dimer, further doubling of concentration leads to tetramers and at 1.5 mg/mL hexamers are observed. In contrast to the short construct from Jehle used for the assignment of the core domain (ABC64-152; αB10.1) that was found only as dimers, in the case of the work of Laganowsky also the C-terminal IXI-motif was present. The detection of tetramers and hexamers in consequence implicates involvement of the C-terminal domain in oligomerization, and in reverse conclusion the N-terminus must be responsible for formation of higher order oligomeric ensembles. In fact, ABC68-162 forms a runaway domain upon crystallization, a polymeric chain built of core domain dimers, linked via binding of the C-terminal IXI-motif to the hydrophobic groove of another subunit. The sequence stretch around the motif is palindromic, namely 156 – ERTIPITRE – 164, and Laganowsky observed bidirectional binding to the groove formed by β-strands 4 and 8, which is facilitated through a flexible hinge loop connecting the core with the C-terminus.

The N-terminus was partially assigned in the solid-state NMR investigations but turned out to be very heterogeneous. Only for a few residues multiple resonances were observed but an overall assignment was not possible[67,76]. Chemical shift analysis with TALOS conducted by Jehle *et al.* predicted two β-strands for Leu44-Tyr48 and Ser59-Thr63 and distance restraints suggest antiparallel arrangement. Furthermore, although only with low confidence, the secondary structure prediction resulted in two helices for the residues Phe14-Phe17 and Phe27-Leu32[76]. Nevertheless, the sparse NMR-assignments did not allow structure calculation. A sequence similarity search was conducted and resulted in similarity with residues 12-62 of xylan esterase from *Thermotoga maritima* (PDB 1vlq, 47% similarity), residues 5-27 of 2'-specific/double-stranded RNA-activated interferon-induced antiviral protein 2'5'-oligoadenylate synthetase (PDB 1px5, 65% similarity) and residues 25-48 of methyltransferase-fold protein from *Erwinia carotovora atroseptice* (PDB 2p7h, 54% similarity). Those matches contain either β-strands or α-helices, which align with one or more of the predicted secondary structures by TALOS. Heterogeneity in NMR-signals can be interpreted in a way that N-terminus does not exist as a single structure at a time. Jehle *et al.* presented a 24mer-model (**Figure 6 B**, PDB-entry 3J07), based on the modelling of the triangular ACD in the electron density map of αB-crystallin[75]. The 24mer included modelling of the N-terminus according to the alignment with the esterase and the methyltranferase[76], leading to a fragment (residues 1–20) that contains helix α1 and a fragment containing α2, β1, and β2 (residues 21–65). Furthermore, the model comprises a C-terminus, which is in parts very flexible. The particle has an inner cavity of 4 nm in accordance to observations from SAXS and EM-studies. The term 'hierarchy of interactions' resulting in oligomerization was introduced[76]. Here, the homo-dimer serves as the main building block with the two β6+7 strands aligned in an antiparallel fashion. The second level of interaction is provided by the C-terminal IXI-motif binding to a hydrophobic groove[64,68], formed by β4 and β8 of an adjacent monomer, and thereby resulting in a triangular arrangement of the ACD-homo-dimers. The third level is the interaction of the N-terminal β2 forming inter- and/or intramolecular contacts with β3 and β4 and thereby building higher order multimers.

Braun *et al.*[77] published in the same year a model for a 24mer (**Figure 6 C**, PDB-entry 2YGD) which is based on Cryo-EM, the NMR-results on the ACD of Jehle and molecular modelling for the N-terminus. The modelling resulted in α-helices that can be observed as rods in the Cryo-EM map of particles corresponding to the 24mer. The particles possess a large inner cavity and 3-fold symmetry axes arranged in a tetrahedral manner as in the model of Jehle.
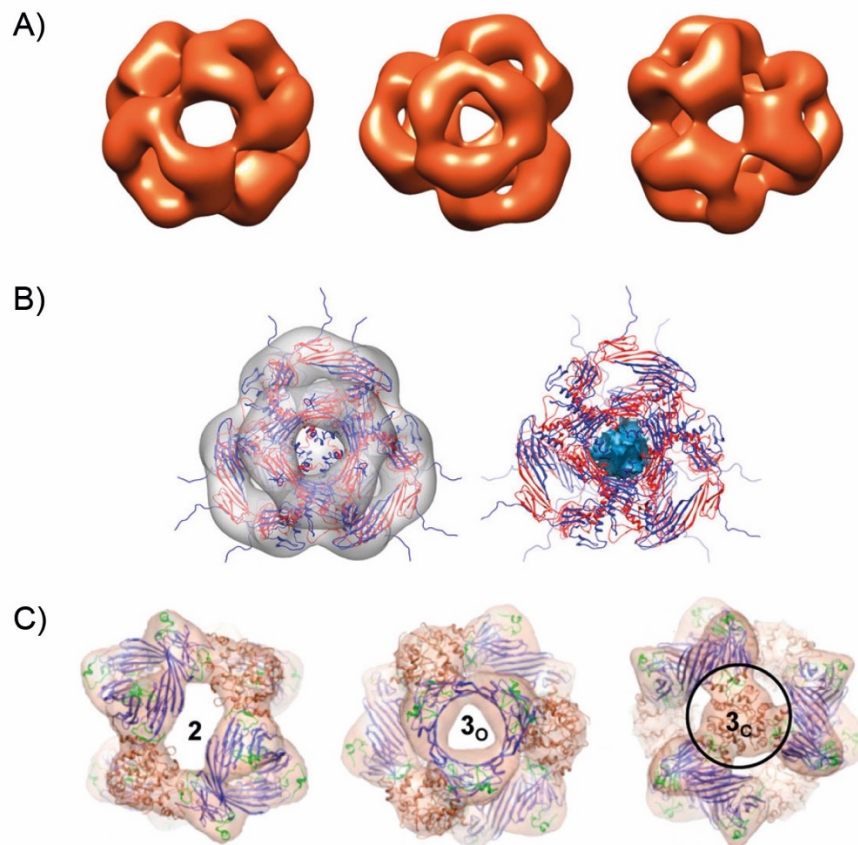
Figure 6: A) EM-structure of αB-crystallin by Peschek *et al.*[75] illustrating the 24mer as spherical hollow particle with tetrahedral symmetry. B) 24mer model of full length αB by Jehle *et al.*[76] based on fitting of the solid-state NMR structure into the EM electron density map and modelling of the N-terminus according to sequence alignment. The flexible C-termini are protruding from the EM density. Marked in blue is the inner cavity with a diameter of 4 Å. C) 24mer model of full length αB-crystallin by Braun *et al.*[77] based on the NMR-structure for the core domain, Cryo-EM and modelling of the N-terminus into the electron density map. Numbers indicate the symmetry of the respective axes through the spherical particle.

The ACD dimers are built out of asymmetric monomers in a bent and an extended form (**Figure 7**), which interact on the one hand at their dimer interfaces via the antiparallel arrangement of the long β-strands, and on the other hand via interactions of all domains. The N-termini of all monomers cluster at a hexameric

interface containing three pairs of N-termini of the extended and bent monomers arranged in a three-fold symmetry.
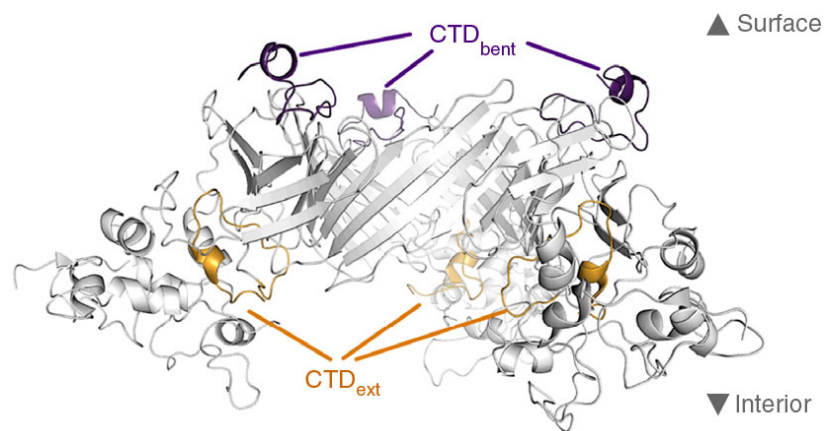
Figure 7: Visualization of bent (CTD$_{bent}$, yellow) and extended (CTD$_{ext}$, purple) C-termini within the hexameric barrel-like structure within the oligomer-model from Weinkauf, see also figure 6 C[77]. Figure taken from [74].

The C-terminus of αB-crystallin is a paradigm on its own. Carver *et al.* revealed via early solution NMR-measurements its flexibility, as for the last 10 residues, signals were observed in TOCSY and ROESY spectra[59], whereas Jehle *et al.* found that the C-terminal IXI-motif is bound to the hydrophobic groove under solid-state NMR conditions[67]. Nonetheless, results from Baldwin et al. indicated that the C-terminal motif is mostly free in solution (to a fraction of 98%) and is only transiently bound to the groove[73,78]. The results were interpreted as the monomers within the oligomeric ensemble residing all in the same conformation, contradicting the asymmetric dimer presented in both 24mer models. The measurements furthermore revealed inter-dimer contacts between the C-terminus of one monomer with the core-domain of another and this interaction was interpreted to derive from residues upstream the IPI-motif. Mainz *et al.* could partially reconcile the conflicting results by confirming the flexible nature of the C-terminus and assigning the residues upstream from Ser153 in HSQC spectra. For some C-terminal residues an additional minor set of NH-signals was observed. Nevertheless, the bound form of the IPI-motif gave rise to signals in Frosty-MAS solid-state NMR spectra of wild-type αB-crystallin solutions[74] leading to the conclusion that the C-terminus must exist in at least two different conformations.

The solid-state NMR results on the full-length construct reveal two signals for H104, which is not the case for the short construct comprising the dimeric ACD. This finding speaks for an asymmetric arranged dimer in the full-length construct, with the asymmetry being mediated by the termini of the protein.

1.1.2 Homologues of αB-crystallin: similarities and differences

The classification of the two α-crystallins as small heat shock proteins was based on homology[79,80] and the α-crystallin domain emerged as the common feature of all sHSPs[13]. Interestingly, even though the protein family members are sequentially rather different, the architecture of the core domain is highly conserved[68] and known structures overlay almost perfectly, implying high structural preservation[81]. Poulain *et al.* systematically analyzed the composition of sHSPs and defined everything between β2 and β9 as ACD, which has a length of 90±10 residues, differing mainly in the length of the loops between the β-sheets for different species. Multiple structure and sequence alignments lead to 3787 sequences owning an α-crystallin domain. The N-terminal region before β2 varies greatly in length (53±35 residues) and sequence and is usually rather hydrophobic. These features hindered systematic investigations of the N-termini among the sHSPs. The residues following the ACD belong to the C-terminal domain, which can be further divided into a C-terminal anchoring module (CAM) with the highly conserved IXI-motif and the C-terminal extension. The CAM makes contacts to the ACD with the hydrophobic motif. If the homology-search is expanded from isoleucine and valine to also include chemically similar leucine residues, the conservation among small heat shock proteins raises from 90% for a V/I – X – I/V-motif to 96% as L/V/I – X – I/V/L-motif.

So far, ten human sHSPs have been recognized[82], HSPB1 (Hsp27, Hsp28), HSPB2 (MKBP), HSPB3 (Hsp17), HSPB4 (CRYAA, αA-crystallin), HSPB5 (CRYAB, αB-crystallin), HSPB6, HSPB7 (CVHSP), HSPB8 (CRYAC, E2IG1, Hsp22), HSPB9 (CT51) and HSPB10 (ODFP1), with αB-crystallin being the most ubiquitous among them. Sequentially closest to αB, with a sequence identity of 51%, is the homologue αA-crystallin. In addition to αA, αB-crystallin is able to form hetero-oligomeric complexes with HSPB1[83], HSPB6[84] and HSPB8[84,85].

1.1.3 Function and posttranslational modifications of αB-crystallin

The chaperones act as a first line of defense under conditions of cellular stress, like elevated temperatures, drop in pH[86–88], UV-radiation[89–91] or oxidative stress[92] and misfolded proteins are bound in an ATP-independent manner. Refolding can take place when the stress situation is overcome in interaction with other proteins like HSP70[7,8,93].

αB-crystallin was first described as a chaperone when Horwitz tested the capability of α-crystallin to prevent the temperature dependent aggregation of alcohol dehydrogenase and βL-crystallin[44]. He furthermore showed that the α-crystallins are responsible for the maintenance of the refractive index in the vertebrate eye-lens. He demonstrated the double role in the lens, as the crystallins have an effect on light scattering and therefore on the refractive properties of the medium and act as chaperones for other eye-lens proteins[94]. Since the protein-turnover in this particular tissue is close to zero, the high solubility, the longevity and transparency-properties are all essential to the crystallins and a reliable interplay between the eye-lens proteins is indispensable. Still, with ongoing aging, the lens proteins can start to aggregate and thereby cataract is developed. This disease comprises a clouding of the eye lens and is the major reason for acquired blindness worldwide[19].

Over the years, the list of known substrates for αB-crystallin grew very long. Examples are the citrate synthetase and α-glucosidase[7], insulin[95], α-Lactalbumin[62], γ-crystallin[19], lysozyme, α-synuclein, Aβ and Tau[25,27,37,48,74,86,96], titin, desmin and actin *in vivo*[97,98]. Upregulation of αB is further correlated with certain cancer types[33–35] and apoptosis, the process of programmed cell death[14].

The remarkable substrate palette of the ubiquitous αB-crystallin is possible due to the fact that the protein has more than one client binding site[99]. Short constructs usually remained functional at least towards certain targets, like the ACD[66,71] or N-terminal residues of the protein[19,95,100,101]. There is no evidence that the cavity of the particles would serve as binding site. On the contrary, the particle size grows upon binding of client protein[62], implying that binding happens at the surface.

Several active sites have been identified in the sequence: the first part of the N-terminus[102], a phenylalanine rich region in the N-terminus 22 – RLFDQFF – 28[19], the

sequence stretches $41 - 58^{101}$ and $54 - 60^{95}$ in the N-terminus, the regions corresponding to β-strands 3[103], 4 and 8 in the core domain[74,86], the ACD as such[71,74] and the C-terminal region[19,46,100,101,104].

Activation of the chaperone through pH-drop was interpreted by Jehle[68] as loosening of the C-terminus resulting in larger particles. The release of the C-terminal IXI-motif from the hydrophobic groove in order to make it accessible for client proteins implies an autoinhibitory function. Jehle furthermore recognized the pH-sensitivity of the dimeric building block[67], a finding which was affirmed by Rajagopal *et al.*[72]. This pH sensitivity was derived from NMR changes in the histidine-rich ACD with His104 being an important pH-trigger for dimer dissociation upon pH-drop below 7.0. The oligomerization of chaperones can regulate their activity[105–107], but these regulation-mechanisms are very complex. Self-regulation with respect to features involved in oligomerization is believed to play a role[68]. According to some studies, smaller oligomers seem to be more active[105,106], while according to others the undisturbed oligomeric equilibrium appears as the more active ensemble[107]. Regulation of the chaperone furthermore takes place through phosphorylation, which is the main posttranslational modification of the protein and the respective sites Ser19, Ser45 and Ser59 are all situated in the N-terminus[108]. Phosphorylation of these serines alters the oligomer-sizes[28,50,109,110] and is found to play a role in chaperone-activation[111]. In the Weinkauf-model, the physphorylation-sites are situated in the hexamer interface of the N-termini and exhibit different accessibility[77]. In brains of patients suffering from Alexander's disease, S59-phophorylated αB-crystallin was found to be coprecipitated with the Rosenthal fibers whereas nonmodified protein stayed soluble[29]. The phosphorylation-mimicking mutant 3D-αB is found to be less active as a chaperone towards luciferase under heat shock conditions compared to the wild-type protein[28]. Although some studies suggest that there is no difference in chaperone capacity between phosphorylated state and wild-type[112–114]there are many other studies that contradict this finding. In case of the 3E-phosphorylation-mimicking mutant, the chaperone activity towards malate dehydrogenase was higher compared to the wild-type form of αB-crystallin[110]. Phosphorylation and formation of mixed oligomers *in vivo* is therefore believed to play a key role in the modulation of the chaperone activity[115–117] and mixed oligomers might bear the ability for further

protein targets not known[118]. Furthermore, in many *in vitro* experimental setups to study αB-crystallin, low concentrated solutions of the protein are used. However, under native conditions within the cell the medium is much more crowded and in investigations mimicking these circumstances, αB-crystallin tends to exist in smaller sized particles[119]. It is well possible that the smaller species are more active as chaperones because of the altered surface-volume ratio, which allows more active sites to be exposed.

1.2 Nuclear magnetic resonance spectroscopy

For the complete understanding of the functional properties of a protein, it is indispensable to know its structure. In order to assess the structure of a protein, a range of different structural biology methods can be used, including X-ray crystallography, cryo-electron microscopy and nuclear magnetic resonance-spectroscopy. About 89% of all structures in the protein database (PDB) are solved by X-ray Crystallography, rendering this method the most frequently used for structure determination. The prerequisite for X-ray crystallography is the successful formation of high-quality single crystals of proteins, which in turn requires relatively large amounts of the recombinant protein. As the formation of a protein-crystal comes along with an environment, which can be quite different from the native environment of the protein, the reliability of structures assessed by X-ray crystallography can be challenged in certain cases. The difficult assessment is particularly true for many membrane-proteins, large protein complexes or heterogeneous proteins. As alternative method, cryo-EM has lately undergone large progress in the technical development and is more and more emerging as a method for structure determination. However, it has, together with X-ray, the big disadvantage that it is not suitable for the investigations of dynamics. NMR-spectroscopy is a well-established method in structural biology. Depending on sample requirements with regard to solubility and natural environment, both solution or solid-state NMR can often be applied to provide, often complementary, structural and structure-dynamical information. One of the biggest advantages of NMR is that it can be used to study proteins in a native-like environment at ambient

temperature, therefore providing full access to study both structure and dynamics of proteins.

1.2.1 Basic principles of NMR-spectroscopy

This chapter aims to explain the general concepts of NMR-spectroscopy and specifically the use of it for investigations of large biomolecules. For a deeper understanding and a more detailed description of the method the reader is kindly referred to the literature[120–124].

The observations in NMR-spectroscopy rely on the effect of an external magnetic field B_0 on the nuclear spins[125–128]. Some commonly used nuclei for NMR-spectroscopy on biological samples are listed in **Table 1.**

Table 1: Spin-number, gyromagnetic ratio γ and natural abundance of selected NMR-relevant isotopes[129].

Nucleus	spin [I]	γ [10^7T^{-1}s^{-1}]	natural abundance [%]
^{1}H	½	26.75	99.985
^{2}H	1	4.11	0.015
^{13}C	½	6.73	1.108
^{15}N	½	-2.71	0.37
^{19}F	½	25.18	100.0
^{31}P	½	10.84	100.0

For a spin of ½, the spin can be parallel or antiparallel to the direction of the external magnetic field B_0, giving rise to two states (or levels) with different energy. This effect is known as the Zeeman effect, and is responsible for the energy difference between the parallel (or spin up, α) and antiparallel (spin down, β) state. These levels are referred to as the Zeeman-levels (**Figure 8**).
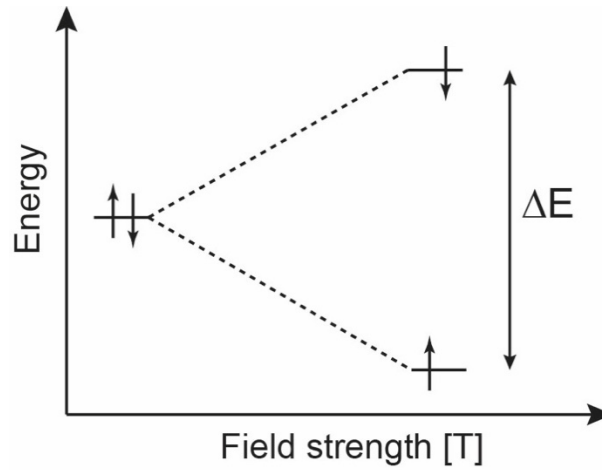
Figure 8: Schematic representation of the Zeeman energy levels of spins in a static magnetic field B_0. ΔE is defined in Equation 1.

Quantum mechanically, the energy difference between the states is proportional to the gyromagnetic ratio γ and the magnetic field B_0,

$$\Delta E = h\nu = \hbar\gamma B_0 \qquad \text{Eq. 1}$$

with the reduced Planck-constant $\hbar$:

$$\hbar = h \,/\, 2\pi \qquad \text{Eq. 2}$$

The energy difference is hence larger for nuclei with high gyromagnetic ratios or in larger external B_0 fields. Assuming Boltzmann distribution for the amount of spins with specific orientation, this corresponds to:

$$\frac{n_\beta}{n_\alpha} = e^{-\Delta E/kT} = e^{-\hbar\gamma B_0/kT} \qquad \text{Eq. 3}$$

Here n_β corresponds to the population of the upper energy-level and n_α to that of the lower, k represents the Boltzmann-constant and T the absolute temperature in Kelvin. The difference in energy between the upper and lower level determines the population difference and hence the net magnetic moment of all spins. Since, in turn, the net magnetization determines the sensitivity of this spectroscopic method, the

usage of magnets with high field strengths and investigation of nuclei with high gyromagnetic ratios γ provide favorable NMR conditions.

The spins nutate around the external field B_0 with a frequency that corresponds to the Larmor-frequency ω_0:

$$\omega_0 = \gamma B_0 \qquad \text{Eq. 4}$$

However, the actual nutation frequency is slightly changed by the physical interactions with other spins and the surrounding electrons. The Hamilton operator quantum mechanically describes these interactions for a given spin I and has the following form:

$$H = H_Z + H_{RF} + H_{CS} + H_J + H_Q + H_D \qquad \text{Eq. 5}$$

Each term in the Hamilton operator describes a certain interaction. The H_Z term describes the Zeeman interaction, H_{RF} interactions with external rf-fields ("NMR pulses"), H_{CS} represents the chemical shift Hamiltonian, H_J takes into account the J-couplings between the spins, H_Q describes the quadrupolar interaction and finally, H_D contains the dipolar interactions between the spins.

The Zeeman effect is given by:

$$H_Z = -\gamma_I \hbar I_z B_0 \qquad \text{Eq. 6}$$

The two spin states α and β correspond to the so-called eigen-states of the operator I_z and represent a state in which the spin is either 100% up (α) or 100% down (β). Any spin state corresponding to a general, non-longitudinal orientation can be quantum-mechanically written as a linear combination of the states α and β.

The Hamiltonian H_{RF} contains the time-dependent magnetic interactions of the spin with an external rf-field, in this form a rf-pulse in the x-direction.

$$H_{RF} = -\gamma_I \hbar I_x B_I \qquad\qquad \text{Eq. 7}$$

In order to induce transitions between the spin states through the rf-pulse, the rf frequency has to be close to the larmor-frequency of the given nucleus, as the strength of the rf-field is comparably small with respect to the B_0 field.

H_{CS} is the Hamiltonian for the chemical shifts and describes the differences in the precession frequencies caused by the local surrounding of the spin. According to Lenz' law, local fields are created by the electrons surrounding a nucleus, which are opposite to the B_0 field. Depending on their shape and density of the electron cloud surrounding the observed spin I, the observed chemical shift varies.

$$H_{CS} = \gamma_I \hbar I_z \sigma_{zz} B_0 \qquad\qquad \text{Eq. 8}$$

σ_{zz} is a component of the shielding tensor σ, which is in most cases not spherically symmetric and hence orientation-dependent. The Hamiltonian H_{CS} for the chemical shift is thereby dependent on the orientation of the molecule within the magnetic field B_0, which makes this interaction anisotropic. The orientation-dependency, however, averages out through the rapid molecular tumbling of molecules in solution. What remains is the isotropic part of the shielding tensor, σ_{iso}, which can be expressed as $(\sigma_{xx} + \sigma_{yy} + \sigma_{zz})/3$, which is not orientation dependent.

J-couplings between spins I and S are mediated through electrons in bonds between the spins and comparably weak.

$$H_J = J\vec{I}\vec{S} \qquad\qquad \text{Eq. 9}$$

J describes the coupling constant and usually has the magnitude of several tens of Hz.

The Hamiltonian H_Q for the quadrupolar interaction is zero for spins ½ and will therefore not considered here since the work in this Thesis only involves spins ½.

The dipolar coupling Hamiltonian H_D is described by

$$H_D = -d(3cos^2\theta_{IS} - 1)I_Z S_Z \qquad \text{Eq. 10}$$

with

$$d = \frac{\left(\frac{\mu_0}{4\pi}\right)\hbar\gamma_I\gamma_S}{r_{IS}^{\ 3}} \qquad \text{Eq. 11}$$

where θ_{IS} describes the angle between the vector connecting the spins I and S and the magnetic field B_0 and r_{IS} is the distance between the two spins.

The dipolar coupling belongs to the strongest influence of the spins on each other. It can be distinguished between homonuclear and heteronuclear dipolar couplings, e.g. ^{1}H-^{1}H and ^{1}H-^{13}C. Dipolar couplings are anisotropic due to the dependency on θ_{IS}. Due to the fast intrinsic molecular tumbling in solution, the dipolar couplings are normally averaged out and can be neglected in solution NMR applications.

1.2.2 Relaxation

When spins experience a radio frequency (rf) pulse which is in general close to the Larmor-frequency ω_0 and perpendicular to the B_0-field, spin populations are changed. When the net z-magnetization is rotated by such a pulse into the orthogonal xy-plane, it is called a 90° pulse. Following the pulse, the two energy states (α and β) are equally populated and the spins will precess in the transversal plane with the Larmor-frequency ω_0. Differences in local fields by for instance chemical shielding (Eq. 8) results in differences in the precession-frequencies. Hence, the precession frequencies will be dispersed giving rise to the so-called chemical shifts, which reflects the sum of all local interactions and shielding and thus reflects the chemical environment of the atom. Several factors influence the experienced chemical environment of a spin, including covalently bound atoms, H-bridges, van-der-Waals interactions, influence of the solvent and pH. Chemical shifts are usually displayed as field-independent values in *parts per million* (ppm). The electronegativity of an atom is a crucial factor for the *primary* chemical shift distribution. Typically, certain functional groups resonate in certain frequency-ranges, for example carbonyls in the range of 170-190 ppm and methyl-groups in the one-digit ppm area[130]. The specific

three-dimensional environment determines the secondary chemical shift, which contains information on the secondary structure of the respective atom[131].

Two different processes, called the T_1- and T_2-relaxation, drive the subsequent relaxation to the state of thermodynamic equilibrium. The longitudinal T_1-relaxation describes the exchange of magnetization of the spin with the lattice through heat exchange towards a normal Boltzmann-distribution which was disturbed by the rf-pulse. The T_2-relaxation derives from interactions with surrounding nuclei representing small local fields and causes a loss of transverse magnetization in the *xy*-plane. The evolution and decay of magnetization is recorded as 'free induction decay' (FID). The FID is converted from the time dependent to the frequency dependent domain by Fourier transformation, resulting in NMR-spectra[132].

1.2.3 Magic-angle-spinning solid-state NMR

For solids or for large protein-complexes in solution the required fast molecular tumbling of the molecules that provides sharp lines in NMR applications is absent. For the investigation of such big complexes in the solid-state, the concept of magic angle spinning can be applied[133].

The dipolar Hamiltonian is given in Eq. 10. Due to the lacking molecular tumbling in a solid it is not averaged to zero has to be considered. It contains θ_{IS}, the angle between the vector connecting the two spins *I* and *S* and the *z*-axis. The term $(3\cos^2\theta_{IS}-1)$ becomes zero at the angle of 54.74°, the so-called magic angle, which corresponds to the angle of the body-diagonal of a cube with the *z*-axis. When the sample is spun in a solid-state NMR rotor around this particular axis, the time average of θ_{IS} for each spin pair is the magic angle, hence this averages out the dipolar coupling and the anisotropic interactions between the spins[134]. Recent developments in the probe design for solid-state NMR applications facilitates nowadays spinning speeds up to 100 kHz[135,136].

1.2.4 Cross polarization

At moderate spinning speeds, the detection of high γ nuclei is not conducive because the homonuclear dipolar interactions are not cancelled out properly and broaden the

NMR lines. For instance, although protons are the most abundant atoms in biomolecules, proton detection at spinning frequencies <20 kHz does generally not result in reasonable line widths. Nuclei with low gyromagnetic ratios, on the other hand, give rise to narrower NMR lines but are relatively insensitive for NMR detection. However, it is possible to make use of the high γ of the protons by transferring the high polarization from protons to low γ nuclei, like carbon or nitrogen. This process is called cross polarization (CP) and helps to overcome the insensitivity of low γ nuclei[137]. For the transfer, the Hartmann-Hahn-condition[138] has to be fulfilled (Eq. 12).

$$\omega_{1H} = \gamma_{1H} H_{1H} = \gamma_{13C} H_{13C} = \omega_{13C} \qquad \text{Eq. 12}$$

Eq. 12 is valid for a static sample. Eq. 13 describes the Hartmann-Hahn-condition for a spinning sample with n = +1, +2.

$$|\omega_{1H} - \omega_{13C}| = n\omega_r \qquad \text{Eq. 13}$$

The protons are excited, usually with an 90° pulse, after which spin lock pulses (B_1) for ^{1}H and ^{13}C are applied in the transversal plane, effectively creating a splitting of the energy levels of the two spins. When the energy-differences caused by the spin-lock pulses match each other, magnetization can be transferred with conservation of energy from the high γ nuclei to the low γ nuclei. To ensure that the narrow matching conditions are fulfilled, usually a ramp on one of the two spin-lock pulses is applied[139]. The standard pulse-program for the ^{1}H excitation, CP and ^{13}C/^{15}N detection with decoupling of protons is shown in **Figure 9**.
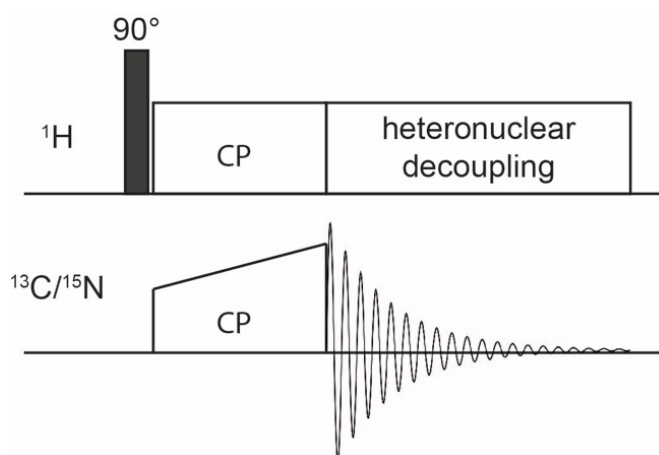
Figure 9: Schematic pulse sequence for a CP-based solid-state NMR experiment. The protons are excited with a 90° pulse and the magnetization transferred to the heteronucleus through a cross polarization step. ^{1}H decoupling takes place during the acquisition of the heteronucleus.

Finally, for a narrower linewidth of the detected (low γ) nuclei, heteronuclear decoupling on protons is usually applied. The cross polarization comes along with the advantage of a faster repetition rate of the experiment as the high γ proton nuclei commonly have a shorter T_1-time and therefore a shorter interscan delay can be used. A common MAS solid-state NMR experiment is to correlate the ^{13}C chemical shifts through usage of proton-driven spin diffusion (PDSD)[140,141]. This method relies on dipolar through-space interactions of nuclei where magnetization is transferred in a diffusion like manner. In solution NMR spectroscopy, magnetization transfer is often based on through-bond *J*-couplings called 'insensitive nuclei enhanced by polarization transfer' (INEPT)[142]. It is possible though to use an INEPT-approach in MAS NMR as well, resulting in spectra displaying the flexible parts of the solid protein.

The backbone assignment of proteins is achieved by the application of triple resonance experiments like HNCACB and HNCOCACB. In this case, the use of CP is extended to selectively transfer between pairs of low γ nuclei such as nitrogens and carbons (N-Cα or N-CO) for backbone assignments in large biomolecules[143]. With this type of experiments the resonances of Cα and Cβ of an amino acid can be linked to the NH-correlation of the same residue *i* (NCA) and the preceding residue *i*-1 (NCO) in the chain.

1.3 Scope of the thesis

The remarkable palette of substrate proteins, the number of disease-states in which upregulation occurs and the variation in substrate affinity in dependency of posttranslational modifications, all indicate an important role of αB-crystallin in the homeostasis of living cells. αB is a heterogeneous protein whose inherent polydisperse nature is most likely essential for its function, in a mechanism that is poorly understood which renders it an interesting target protein for structural biology research. Although the description of structural features of the protein is constantly being refined, so far, there are only models available to describe its quaternary structure and an atomic structure for the full-length protein stays elusive. Being essential for both substrate recognition[19,95,101] and formation of higher order oligomers[70], the role and structural features of the N-terminus remains enigmatic. It is the most heterogeneous part of the protein and its characterization is difficult. In this work we aimed to gain a deeper understanding of the role of the N-terminus and the two terminal IXI-motifs, 3 – IAI – 5 and 159 – IPI – 161, by making use of different NMR-techniques for their investigation. The C-terminal motif is known for binding to a hydrophobic groove formed by the strands β4 and β8[68,70]. The N-terminus similarly contains an IXI-motif at its very beginning, 3-IAI-5. So far, the N-terminal domain is not well described as it withdrew itself from standard investigations methods.

As αB-crystallin is one of the most soluble proteins in the human organism, intuitively solution NMR could well serve as a method for its investigation. However, because of the formation of large oligomeric assemblies, it is not possible to resolve the full-length protein in solution NMR-spectra. Therefore, we use a hybrid approach of solution and MAS solid-state NMR to investigate the complex system. We use both IXI-motifs as anchor points in our investigations by comparing the wild-type protein to selected mutants of the respective motifs. The apparently conflicting results about the CTD motivated us to determine the amount of bound and unbound C-terminal IXI-motif, which we tried to assess with NMR-measurements that allow a direct quantification. With our findings we aim to better characterize the role and the function of both, the CTD and NTD in the heterogeneous ensemble.

2 Materials and methods

2.1 Recombinant protein expression and isotope labelling

The recombinant protein expression was carried out according to the established protocol used in previous works on the αB-system[67]. All wild-type and mutant samples were prepared by **Kristina Rehbein**, **Natalja Erdmann** and **Martina Leidert** under the supervision of **Dr. Anne Diehl**. The work included the site-directed mutagenesis as well as the protein expression and the protein purification.

2.1.1 Recombinant protein expression and purification

E.coli BL21 Star (DE3) cells were used for recombinant protein expression. The cells were grown on M9 minimal media and incubated until reaching a density of 0.6 at 600 nm at 37 °C and then transferred to a 22 °C shaker with subsequent protein expression induction by addition of 1 mM IPTG. Harvesting was done after 12 – 16 h. The lysis and centrifugation steps were done according to a protocol by Feil [66]. The fractions containing protein are purified over a *DEAE-Sepharose* anion-exchange column; washing of the column was carried out with a series of step gradients between Tris-buffer A (20 mM Tris, pH 8) and Tris-buffer B (20 mM Tris, 1 M NaCl, pH 8). After another desalination step (*Sephadex25* column) >90% purity was achieved via anion-exchange column *MonoQ HR10/10* (Pharmacia). Gel filtration over a *Sephadex75* column yielded in >98% protein purity. After dialysis for buffer exchange to sodium phosphate buffer (20 mM Na_2PO_4, 50 mM NaCl, pH 7.8) the resulting protein-material was either used for immediate NMR-sample preparation or lyophilized for storage.

2.1.2 Isotope labelling

For the uniform [^{13}C, ^{15}N]-labelling of the proteins, cells were grown on M9 minimal media with ^{15}NH$_4$Cl (1 g/L) as the only nitrogen source and ^{13}C$_6$-glucose (2 g/L) as the only carbon source.

To facilitate assignments in the 1D-spectra the method of sparse-labelling with 2-^{13}C glycerol (2G) was applied[144–146]. The sparse labelling results in narrower linewidths,

decrease of ^{13}C-^{13}C scalar couplings and improved spectral resolution as only the ^{13}C-labelled carbons contribute to signals in the spectra.
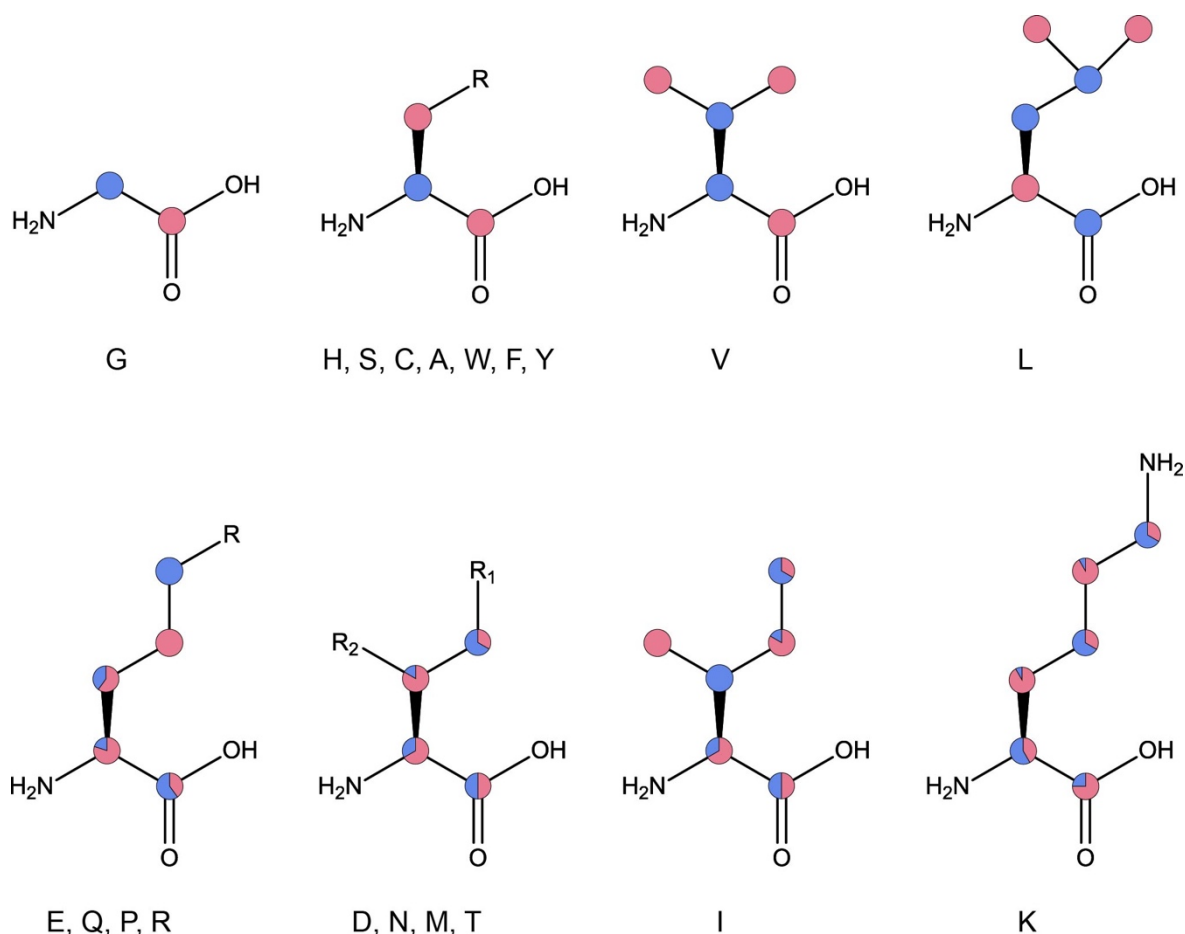

Figure 10: Labelling-scheme for ^{13}C-glycerol sparse labelling. In blue atoms the ^{13}C labelled upon use of 2-^{13}C-glycerol as the carbon source, red labelling-pattern achieved by use of 1,3-^{13}C-glycerol as the only carbon source in recombinant protein-expression. In cases of incomplete isotope labelling the percentage is equivalent to the proportion of the colored amount in the circles.

The glycerol is metabolized in different pathways, including the citric acid circle, the pentose phosphate pathway, and glycolysis. Different synthesis pathways of amino acids result in very specific labelling patterns for the carbons in the various amino acids depending on the primary carbon source. The specific labelling scheme obtained by use of 2-^{13}C-glycerol is shown in **Figure 10**, with the labelled atoms marked blue. Usage of 1,3-^{13}C-glycerol would result in a complementary labelling pattern as pictured by the red color in **Figure 10**.

2.2 Biochemical and biophysical methods and sample preparation

2.2.1 Protein concentration determination

Concentrations of protein-solutions were determined from 2 µL sample volumes via measurement of the absorption of aromatic residues at a wavelength of 280 nm with a *NanoDrop 1000 spectrometer* (Thermo Fisher Scientific). Concentrations were calculated with the extinction coefficient ε (extinction coefficient-values used for the different mutants are listed in **Table 2** and obtained from the *Protparam* database, see 2.4.1).

Table 2: Extinction coefficients ε of αB-crystallin and mutants calculated by ProtParam.

Protein	Extinction coefficient ε (Abs. 1% = 1 g/L)
αB-WT	0,693
αB I3/5V	0,694
αB I159/161V	0,694
αB I3/5/159/161V	0,695
αB A4C	0,692
αB P160C	0,693

2.2.2 Radical synthesis and spin-label reaction

The synthesis of the methanesulfonothioate-derivative of TOTAPOL (1-(TEMPO-4-oxy)-3-(TEMPO-4-amino)propan-2-ol) SB-II-18 (**Figure 21**) for covalent attachment to the cysteine-mutant proteins was done by **Snædís Björgvinsdóttir** at the University of Iceland. The monoradical MTSL (*S*-(1-oxyl-2,2,5,5,-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate) for spin-labelling is commercially available. For the spin-labelling reactions 100 mM stock solutions of the radicals in DMF (Dimethylformamid) were prepared.

The cysteine-mutants of αB-crystallin (A4C and P160C) were buffer exchanged with Zeba™Desalt Spin Columns (2 ml) to sodium phosphate buffer (20 mM Na_2PO_4, 50 mM NaCl, pH 7.8) without DTT (Dithiothreitol) as reducing agent.

In all labelling reactions the radicals were added to the DTT-free protein solutions in 10-fold excess and the mixtures shaken at room temperature for 35 minutes. The excess of spin-label was washed via spin column. The sample preparations from the resulting radical-labelled protein solutions were carried out as described in chapter 2.2.3-2.2.4. Labelled proteins were stored at -80 °C and only thawed when necessary.

2.2.3 Solution-state NMR spectroscopy sample preparation

To all protein-solutions 10 Vol. - % of D_2O were added to the sodium-phosphate buffer (20 mM Na_2PO_4, 50 mM NaCl, pH 7.8). 160 – 200 µL of the resulting sample-solutions were pipetted into 3 mm standard NMR-tubes. Concentrations were determined by nanodrop measurements. For quantitative comparison of spectra, the measurements were carried out with same acquisition parameters and differences in concentrations were taken into account by scaling of the respective spectra.

Table 3: Concentrations of solution-NMR samples determined by Nanodrop measurements; The concentration of the spin-labelled sample was calculated from values obtained before the labelling-reaction.

Sample	Concentration
αB-WT	23.8 mg/mL
αB I3/5V (Nmu)	19.0 mg/mL
αB I159/161V (Cmu)	24.2 mg/mL
αB I3/5/159/161V (Qmu)	25.9 mg/mL
αB A4C	22.7 mg/mL
αB P160C	22.7 mg/mL
αB P160C-T	15 mg/mL

2.2.4 MAS solid-state NMR-spectroscopy sample preparation

Precipitation of protein-samples was done according to the protocol described elsewhere[67]. 100 µL of a PEG 8000 (Polyethylene glycol) stock solution (50 g PEG 8000 in 100 mL H_2O) were pipetted to a solution containing a known amount of protein. For 3.2 mm rotors 8 mg of protein were used, for 4 mm rotors 15 mg of protein were used. The concentrations of the solutions were in the range of 1 to 1,5 mg/mL (20 – 30 mM). The resulting mixture was kept in a glass vessel

which was stored in a larger plastic vessel filled with PEG stock-solution. The edges of the cap were lubricated with grease to prevent vapor diffusion. The vessel was kept for four days at room temperature to assure complete precipitation. The resulting solid was transferred into the NMR rotor via centrifugation. For the wild-type protein and all three valine-mutants 4 mm rotors were used. For the labelled and unlabeled cysteine-mutants 3.2 mm rotors were used. The amount of protein in 4 mm rotors was 15 mg, the final amount in 3.2 mm rotors 7.8 - 8 mg.

2.2.5 Proteolysis with enterokinase

The proteolytic cleavage of the C-terminus, as described elsewhere[74], was carried out by **Kristina Rehbein**. Protein-solutions of 1 mg/mL were incubated with enterokinase at 4 °C for 40 h and the educts and resulting mixtures applied to 12% SDS-PAGE (**Figure 34**).

2.2.6 Dynamic light scattering

Dynamic light scattering measurements were carried out on a *Zen3500 Zetasizer NanoZS* instrument (Malvern Instruments), equipped with a 50 mW laser working at 532 nm. The backscattering was measured at an angle of 173° and the temperature controlled via a peltier-element within the instrument. The scattering Intensity I_S is directly proportional to the hydrodynamic radius R_h of the protein-particle. Estimation of hydrodynamic radiuses R_h was carried out by the software *DTS 5.03* (Malvern Instruments). Due to the dependency of I_S on the size of particles in solution, the scattering for large particles as multimers is more intense. Extraction of the hydrodynamic Diameter D_h is done on artificially intensity-weighed distributions to compensate for the large distribution of particle-sizes in the polydisperse αB-system.

Measurements were carried out at 20 °C after temperature-equilibrating the samples for at least 10 minutes before and 5 minutes between every measurement. Per measurement 30 runs for 5 s were carried out (150 s) and averaged, in total every sample was run 3 times. Data processing was done by the instrument-software. Sample concentrations were 0.15 mM in sodium-phosphate buffer (20 mM Na_2PO_4, 50 mM NaCl, pH 7.8) and placed in a Quartz-cuvette for small volumes (12 µL).

2.2.7 Small-angle X-ray scattering

Small-angle X-ray scattering measurements on the wild-type and mutant αB-crystallin constructs (I3/5V, I159/161V, I3/5/159/161V, A4C and P160C) were carried out by **Dr. Haydyn Mertens** (EMBL Hamburg) on the *P12 SAXS beamline* (Petra III storage ring, DESY, Hamburg). Protein samples were provided at a concentration of 2 mg/mL and volumes of 140-200 μL along with a matching buffer (20 mM Na_2PO_4, 50 mM NaCl) for measurement and subsequent background subtraction. Measurements were carried out on a set of serial dilutions at three different concentrations, 2.0, 1.0 and 0.5 mg/mL. For data analysis and determination of the SAXS parameters *I0*, R_g and D_{max}, the *ATSAS software package*[147] including the indirect Fourier transformation program *GNOM*[148] was used. As basis for evaluation of the data towards an oligomeric distribution in the program *OLIGOMER*[147], models of αB-crystallin as 24mers and sub-assemblies were used, corresponding to PDB entry numbers 3J07 and 2YGD.

2.2.8 Electron paramagnetic resonance-spectroscopy and double electron-electron resonance-spectroscopy

Electron paramagnetic resonance (EPR)-spectroscopy measurements for determination of the labelling efficiency and of the distances between the labels via double electron-electron resonance (DEER)-spectroscopy were carried out by **Prof. Dr. Enrica Bordignon** at the Freie Universität in Berlin-Dahlem.

20 μL of samples of αB-A4C (333 μM) and αB-P160C (233 μM) with MTSL-label (for preparation see 2.2.2) were transferred into 3 mm glass capillary tubes. EPR-measurements were recorded at RT and DEER-measurements acquired under cryogenic conditions at 50 K on a *Bruker ELEXSYS E580Q-AWG* (arbitrary wave generator) dedicated pulse Q-band spectrometer equipped with a 150 W TWT amplifier. The resonator for tubes with an outer diameter of 3 mm is homemade. Background fitting was applied to the primary data for short and long DEER-traces resulted. The resulting curves for secondary data [F(t)] were converted by a model-free Tikhonov regularization to distance distributions[149].

2.3 NMR-spectroscopy, data processing and evaluation

2.3.1 Solution NMR-spectroscopy parameter

Solution NMR-spectroscopy measurements were carried out on a Bruker Avance 600 NMR-spectrometer, equipped with a cryogenically cooled probehead, with a magnetic field strength of 14.1 T and a Bruker 750 NMR-spectrometer with a magnetic field strength of 17.6 T (measurements on the 750 MHz machine marked with * in **Table 4**). Shimming was carried out automatically before every measurement. For the ^{13}C direct excitation experiments the WT was measured at 280, 290, 300 and 310 K, the valine mutants at 290, 300 and 310 K, and the 2-^{13}C-glycerol labelled WT-sample at 290 K. For all multidimensional measurements the temperature was set to 300 K (T_{real}). Recycle delays were set at 4.7 s for HSQCs, at 2 s for CC-TOCSY-experiments and at 1.3 s for all other 2D and 3D spectra. ^{13}C direct excitation experiments were run with 30 s recycle delay to assure complete relaxation between the scans. Parameters such as offset and pulses were, if necessary, optimized for best signals to noise ratio. For sample concentrations see **Table 3**.

Table 4: Experimental parameters for solution NMR-measurements

Experiment	Sample(s)	NS	Coefficient points	Spectral Widths
^{13}C-1D	WT, Nmu, Cmu, Qmu,	2432	262144	331 ppm (^{13}C)
	2G-^{13}C WT	5120		
^{1}H-^{15}N HSQC _sofast	WT	16	1024x256	17 ppm (^{1}H) 50 ppm (^{15}N)
	Cmu	128		
	Nmu, Qmu, P160C-T	256		
^{1}H-^{15}N HSQC	A4C, P160C	32	4096x256	17 ppm (^{1}H) 50 ppm (^{15}N)
^{1}H-^{13}C HMQC	WT	108	1024x256	17 ppm (^{1}H) 79 ppm (^{13}C)
	Nmu	192		
	Cmu	40		
	Qmu	64		
CC-TOCSY	WT, Nmu, Cmu, Qmu	304	8192x128	331 ppm (^{13}C) 83 ppm (^{13}C)
CBCACO	WT	200	1024x128	66 ppm (^{13}C) 66 ppm (^{13}C)
CCCO	WT	232	1024x128	66 ppm (^{13}C) 83 ppm (^{13}C)
HNCA _best	Cmu	16	1024x48 x128	17 ppm (^{1}H) 25 ppm (^{15}N) 33 ppm (^{13}C)
HNCACB _best	Cmu	16	1024x48 x128	17 ppm (^{1}H) 25 ppm (^{15}N) 66 ppm (^{13}C)

HNCOCA_best	Cmu	24	1024x48 x128	17 ppm (^{1}H) 25 ppm (^{15}N) 33 ppm (^{13}C)
HNCOCACB_best	Cmu	24	1024x48 x128	17 ppm (^{1}H) 25 ppm (^{15}N) 66 ppm (^{13}C)
HNCA*	P160C	16	1024x74 x128	17 ppm (^{1}H) 30 ppm (^{15}N) 33 ppm (^{13}C)
HNCACB*	P160C	16	1024x74 x128	17 ppm (^{1}H) 30 ppm (^{15}N) 66 ppm (^{13}C)
HNCO*	P160C	8	1024x96 x128	17 ppm (^{1}H) 30 ppm (^{15}N) 17 ppm (^{13}C)
HNCOCACB*	P160C	16	1024x74 x128	17 ppm (^{1}H) 30 ppm (^{15}N) 66 ppm (^{13}C)
TROSY*	P160C	64	1024x256	17 ppm (^{1}H) 30 ppm (^{15}N)

Referencing of solution-NMR spectra was carried out according to the assignment for αB-crystallin from Andi Mainz deposited in the BMRB with the entry number 26640.

^{13}C direct-excitation spectra were processed with line broadening (LB)-value 5 and automatically baseline-corrected. Subtracted spectra were manually baseline corrected in the area from 100 to 20 ppm and subsequently integrated within the defined regions (**Table 9**). The multi-dimensional spectra were processed with a square sine-bell apodization-function with sine bell shift (SSB) values of 2 to 3.

2.3.2 MAS NMR-spectroscopy parameter

MAS solid-state NMR-spectroscopy measurements were carried out on a widebore Bruker Avance 700 NMR-spectrometer (Bruker Biospin) with a magnetic field strength of 16.4 T with a triple-resonance MAS probe-head for 4 mm and an E-free triple-resonance probe head for 3.2 mm rotors. The magic angle was adjusted with KBr and shimming to linewidths < 5Hz measured with adamantane. MAS-frequencies were set to 12.5 kHz for measurements with 4 mm rotors and 15 kHz in case of 3.2 mm rotors. The temperature was set according to the spectrometer and probe head-specific calibration-curve to result in a real temperature of 291 K. Recycle delays were 4.7 - 5 s for the PDSD-measurements and 2.5 - 5 s for the INEPT-based measurements. PDSD mixing time was set to 15 ms. The number of scans taken for every measurement was 64, except for the HcC-INEPT-COSY for αB P160C-T, where it was 512 and all hCC-INEPT-COSY spectra where 128 scans were taken for every measurement.

Referencing of PDSD spectra was done according to the spectra and assignment of Stefan Jehle whose Sparky-projects were available for this work.

The spectra were processed with a square sine-bell apodization-function with SSB-values of 2 or 3 in the F1 dimension and gaussian function with LB-values of -20.0 to -50.0 and gaussian maximum positions of 0.08 to 0.12 in the F2-dimension. Values were chosen according to best outcome in terms of signal-to-noise and resolution. Usually, when spectra are overlayed in the figures, they are processed in the same way.

Table 5: Experimental parameters for MAS solid-state NMR measurements

Experiment	Sample	Coefficient points	Spectral Widths
PDSD	WT	3326x768	379 ppm (^{13}C) 284 ppm (^{13}C)
	Nmu	1194x768	284 ppm (^{13}C) 284 ppm (^{13}C)
	Cmu	3326x768	285 ppm (^{13}C) 284 ppm (^{13}C)
	Qmu	2658x768	379 ppm (^{13}C) 284 ppm (^{13}C)
	A4C	1394x768	284 ppm (^{13}C) 284 ppm (^{13}C)
	P160C	994x748	284 ppm (^{13}C) 284 ppm (^{13}C)
	P160C-M	994x784	284 ppm (^{13}C) 284 ppm (^{13}C)
	P160C-T	1780x720	255 ppm (^{13}C) 255 ppm (^{13}C)
HC-INEPT	WT	3494x384	284 ppm (^{13}C) 19 ppm (^{1}H)
	Nmu, Cmu, Qmu, P160C	4680x192	381 ppm (^{13}C) 9 ppm (^{1}H)
	A4C	3494x96	284 ppm (^{13}C) 10 ppm (^{1}H)
	P160C-T	4680x92	381 ppm (^{13}C) 9 ppm (^{1}H)
HcC-INEPT-COSY	WT, Nmu, Cmu, Qmu, P160C	3340x128	381 ppm (^{13}C) 9 ppm (^{1}H)
hCC-INEPT-COSY	Cmu, P160C	3340x576	381 ppm (^{13}C) 284 ppm (^{13}C)

2.4.1 Databases and software

Protein sequences were downloaded from the *Uniprot* database at https://www.uniprot.org. Sequence alignments were accomplished by using the *Clustal Omega* program within the *Uniprot* database. The entry codes, names and possible synonyms for the human small heat shock proteins used in the alignment are P04792 (HSPB1, Hsp27, Hsp28), Q16082 (HSPB2, MKPB), Q12988 (HSPB3, Hsp17), P02489 (HSPB4, CRYAA, αA-crystallin), P02511 (HSPB5, CRYAB, αB-crystallin), O14558 (HSPB6), Q9UBY9 (HSPB7, CVHSP), Q9JUY1 (HSPB8, CRYAC, E2IG1, Hsp22), Q9BQS6 (HSPB9, CT51), Q14990 (HSPB10, ODFP).

Protein parameters like isoelectric point pI, extinction coefficient ε and molecular weight M_R were calculated with *ProtParam* available at http://web.expasy.org/protparam/.

Assignment lists for solution-NMR were downloaded from the *Biological Magnetic Resonance Data Bank (BMRB)* at http://www.bmrb.wisc.edu.

Structure models of αB-crystallin 24mers with entry numbers 2YGD and 3J07 were downloaded from the *Protein Data Bank (PDB)* at https://www.rcsb.org

Processing of NMR-spectra was accomplished with *TopSpin 3.5pl7* provided by Bruker Biospin.

Sparky was used for assignments (NMR-assignment software, version 3.115 for Mac OS). The program is available for download at https://www.cgl.ucsf.edu/home/sparky/.

Plots were prepared using the plot-editor in *TopSpin* and postprocessed with *Adobe Illustrator*. Figures were prepared with *Adobe Illustrator* and in case of 24mer models with *PyMol*, available at https://pymol.org/. Distance measurements within the models were carried out in *PyMol*.

3 Results and discussion

3.1 The L/V/I – X – I/V/L – motif in αB-crystallin and human homologues

The C-terminal IXI-motif, or rather the extended L/V/I – X – I/V/L -motif, is highly conserved in all known small heat shock proteins[81] as it is found in 96% of all C-terminal domains of the sHSPs. This finding resulted from the systematical investigation of the CTDs of approximately 40000 heat shock proteins in the protein database, which was possible because of the similar hydrophilicity and length of the C-terminal domains. A similar systematic investigation on the N-termini was not possible because they vary greatly both in hydrophobicity and length.

Taking a closer look at the alignment of the ten human small heat shock proteins (**Figure 11**), it is striking to observe that the conservation of the extended motif in the C-termini is only 60% - which is rather low compared to the results from Poulain *et al.* (96%). The motif is present in HSPB1, HSPB2, αA-crystallin, αB-crystallin and HSPB7 as well as in HSPB9. In the latter, the spacing to the α-crystallin domain is nonexistent as the motif starts directly after the ACD. For all other C-terminal motifs they are situated approximately 10 residues apart from the core-domain.

Interestingly, in HSPB1, HSPB2 and HSPB9, the motif is extended, and has three hydrophobic residues separated by two X as spacers: 179 – ITIPV – 183 in HSPB1, 160 – VYISL – 164 in HSPB2 and 132 – VALAL – 136 in HSPB9. One could speculate that this extended motif L/V/I – X – L/V/I – X – I/V/L can bind to the hydrophobic groove in two different register shifts. A similar feature was observed for the dimer interface of the core domain, where the extended β6+7 strands of two monomers can bind in three different register shifts to each other[70].

Poulain *et al.* were not able to examine the NTD of sHSPs in a similar systematic fashion as it was done for the CTD, due to the high degree of variation in hydrophobicity and length. Nevertheless, a closer look at the relatively small group of human heat shock proteins reveals that an IXI-motif is present in the N-termini of five out of ten human sHSPs, which suggests a remarkable high degree of conservation. It is found in HSPB3, αA-crystallin, αB-crystallin, HSPB6 and HSPB9.

Again, in the case of HSPB9, the spacing to the α-crystallin domain is shorter compared to the human homologues. HSPB6 contains an extended motif 3 – IPVPV – 7 as described for the C-terminus in HSPB1, HSPB2 and HSPB9, whereas HSPB3 not only has a doubled motif but an additional single motif as well (4 – IILRHLIEIPV – 14).

```
HSPB1 MTERRVPFSLLRGPSW-----DPFRDWYPHSRLF---DQA-------FGLPRLPEEWSQWLG-------GSSWPGYVRP 57
HSPB2 MSGRSVPHAHP---------ATAEYEFANPSRLG---EQR-------FGEGLLPEEILT--P-------TLYHGYYVRP 51
HSPB3 MAKIILRHLI----------------EIPVRYQ---EEF-------EARGLEDCRL-------------DHALYALP 38
CRYAA -------MDVTIQHPWFKRTLGP---F-YPSRLF---DQF-------FGEGLFEYDLLPFLS-------STISPYY-RQ 50
CRYAB -------MDIAIHHPWIRRPFFP---FHSPSRLF---DQF-------FGEHLLESDLFPTST-------SLSPFYL-RP 51
HSPB6 -----MEIPVPVQPSWLRRASAPLPGLSAPGRLF---DQR-------FGEGLLEAELAALCP-------TTLAPYYLRA 57
HSPB7 MSHRT-------SSTF----RAERSFHSSSSSSSSSTSSSASRALPAQDPPMEKALS---------MFSDDFGSFMRP 58
HSPB8 MADGQMPFSCHYP-SRL--RRDPFRDSPLSSRLL---DDG-------FGMDPFPDDLTASWPDWALPRLSSAWPGTLRS 66
HSPB9 ---------------MQ---RVGNTFSNESR----------------------VASRCPS--------VGLAERN 27


HSPB1 LPPAAIESPAVAAPAYSRALSRQLSSGVSEIRHTADRWRVSLDVNHFAPDELTVKTKDGVVEITGKHEERQDEHG---Y 133
HSPB2 RAAP--------------AGEGSRAGASELRLSEGKFQAFLDVSHFTPDEVTVRTVDNLLEVSARHPQRLDRHG---F 112
HSPB3 GPT--IVDLR---KTRAAQSPPVDSAAETPPREGKSHFQILLDVVQFLPEDIIIQTFEGWLLIKAQHGTRMDEHG---F 109
CRYAA ------------SLF----RTVLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVQDDFVEIHGKHNERQDDHG---Y 109
CRYAB ------------PSFLR-APSWFDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGDVIEVHGKHEERQDEHG---F 113
HSPB6 ------------PSV--------ALPVAQVPTDPGHFSVLLDVKHFSPEEIAVKVVGEHVEVHARHEERPDEHG---F 112
HSPB7 HSEP------------LAFPARPGGAGNIKTLGDAYEFAVDVRDFSPEDIIVTTSNNHIEVRA---EKLAADG---T 127
HSPB8 GMVPR------GPTA-TARFGVPAEGRTPPPFPGEPWKVCVNVHSFKPEELMVKTKDGYVEVSGKHEEKQQEGG---I 134
HSPB9 RVAT-------MPV---RLLRDSPAAQEDNDHARDGFQMKLDAHGFAPEELVVQVDGQWLMVTGQQQLDVRDPERVSY 95
                                       :  ::.  * *:::  :  .       :  :  .


HSPB1 ISRCFTRKYTLPPGVDPTQVSSSLSPEGTLTVEAPMPKLAT--QSN--EITIPVTFESRAQLGGPEAAKSDETAAK--- 205
HSPB2 VSREFCRTYVLPADVDPWRVRAALSHDGILNLEAPRGGRHLDTEVN--VYISLLPA-------PPDPEEEEEAAIVEP 182
HSPB3 ISRSFTRQYKLPDGVEIKDLSAVLCHDGILVVEVKDPVGTK----------------------------------- 150
CRYAA ISREFHRRYRLPSNVDQSALSCSLSADGMLTFCGPKIQTGLDATHA--ERAIPVSREEKPTSAPSS----------- 173
CRYAB ISREFHRKYRIPADVDPLTITSSLSSDGVLTVNGPRKQV----SGP--ERTIPITREEKPAVTAAPKK---------- 175
HSPB6 VAREFHRRYRLPPGVDPAAVTSALSPEGVLSIQAAPASAQAPPPAA--AK------------------------- 160
HSPB7 VMNTFAHKCQLPEDVDPTSVTSALREDGSLTIRARRHPHTEHVQQ-TFRTEIKI----------------------- 170
HSPB8 VSKNFTKKIQLPAEVDPVTVFASLSPEGLLIIEAPQVPPYSTFGESSFNNELPQDSQE-------------VTCT---- 196
HSPB9 RMSQKVHRKMLPSNLSPTAMTCCLTPSGQLWVRGQCVALALPEAQTGP---------SPRLGSLGSKASNLTR----- 159
             :    :*  :.    :  .  *   .* *  .
```

Figure 11: Sequence alignment of nine out of ten human sHSPs (HSPB1, HSPB2, HSPB3, CRYAA, CRYAB, HSPB6, HSPB7, HSPB8 and HSPB9)[82]. Alignment was done with the *Clustal Omega* program within the Uniprot-database. Completely conserved residues marked with a star, highly conserved residues with colon, residues with medium conservation marked with dot. Sequence stretches comprising I/V/L – X – L/V/I are highlighted in yellow.

In total, there are three human homologues containing a motif in both termini (αA, αB and HSPB9), four homologues with a motif in only one of their termini (HSPB1, HSPB2, HSPB6 and HSPB7) and two lacking the motif completely (HSPB8 and ODFP1). The outer dense fiber protein ODFP1, the tenth human small heat shock protein, is only poorly understood so far. Its function is not known, neither its structure and degree of polydispersity[150]. The termini, especially the C-terminus varies more than usual in its composition compared to homologues of the

superfamily of small heat shock proteins and was therefore not considered in the sequence alignment.

Given the high conservation-rate of the IXI-motif, it is no surprise that the C-terminal motif in αB-crystallin was considered to play an important role in the oligomerization process[67,70]; moreover, it's binding mechanism strongly suggests that it may function as a "gatekeeper" by covering the hydrophobic groove in order to control the chaperoning function. In contrast, the N-terminal motif received little interest in the published data on sHSPs and α-crystallins in particular. Only in case of Tsp36 from the parasitic flatworm *Taenia saginata,* the N-terminal IXI-motif is known to bind to the α-crystallin domain in a similar fashion[151]. But, as it is fairly conserved in the N-terminal position among the human homologues it is tempting to attribute a very similar biological role to it in the oligomerization and auto-regulation process as the C-terminus.

3.2 Wild-type αB-crystallin and selected valine-mutants

For a systematic investigation of the two motifs in αB-crystallin, we carried out NMR experiments with the wild-type and a selected set of valine-mutants of the protein, αB I3/5V (N-terminal mutant, Nmu), αB I159/161V (C-terminal mutant, Cmu) and αB I3/5/159/161V (quadruple mutant, Qmu). NMR-spectroscopy is highly sensitive for small changes at atomic level. Isoleucine side chains provide very characteristic chemical shifts, which enables us to track changes in mutants with comparably subtle modifications compared to the wild-type protein, by systematically replacing specific isoleucines by valines. Taking into account the natural variation of the extended motif as described by Poulain[81], according to which isoleucines and valines are often interchanged, we do not expect a disturbance of the biological function of the termini by exchanging isoleucines with valines. The integrity of the oligomers of mutant αB-crystallin was confirmed by SAXS and DLS measurements. The wild-type and valine-mutants show all similar distribution of oligomeric sizes in DLS and SAXS-measurements. The hydrodynamic diameters D_h measured by DLS vary in the range of the error of the measurement. The values obtained with the two methods are listed in the **Table 6**. To put this in a broader perspective, comparable studies on glycine

and alanine mutants of the C-terminal motif (I159/161G[46] and I159/161A[104]) revealed only minor changes in oligomeric size. In this context it should be noted that a mutation of isoleucine to glycine or alanine afflicts a larger chemical modification than a mutation from isoleucine to valine. Also the tertiary structure of the I159/161A-mutant did not alter significantly according to CD-spectroscopy results[104].

In order to obtain structural information with atomic detail and to study the role of the termini, MAS solid-state NMR-measurements on wild-type αB-crystallin and the three valine-mutants (Nmu, Cmu and Qmu) were carried out on $^{13}C,^{15}N$ labelled protein-samples. The oligomers were precipitated with PEG 8000 and centrifuged into 4 mm rotors. For solution-NMR measurements, protein samples with the same labelling scheme were used at concentrations of 1-1,5 mM (for details of sample-preparation, see 2.2.3-2.2.4.).

Figure 12: Schematic picture of composition of the IXI-motifs and the valine-mutants used in the study. In the top row is the WT with two IXI-motifs, second the Nmu αB I3/5V, third the Cmu αB I159/161V, bottom the Qmu αB I3/5/159/161V.

Table 6: Hydrodynamic Diameters D_h assessed by DLS and SAXS-measurements. Concentration of DLS samples was 0.15 mM and 0.1 mM for SAXS-samples.

Method	WT	Nmu	Cmu	Qmu
DLS	16,41 nm	15,78 nm	17,14 nm	16,01 nm
SAXS	15,54 nm	17,00 nm	17,00 nm	16,00 nm

3.2.1 MAS solid-state NMR of αB-crystallin and valine-mutants

In previous work by Jehle *et al.,* the assignment of solid-state NMR spectra and the subsequent structure calculation of the crystallin domain of αB-crystallin, in combination with SAXS investigations on the oligomeric ensemble, lead to the model

of a 24mer[68]. With use of MAS solid-state NMR, Jehle *et al.* were able to assign the core domain of the protein, as well as the first part of the CTD including the C-terminal IXI-motif, for which they observed contacts to a hydrophobic groove formed by β-strands 4 and 8. The N-terminus and the very end of the C-terminus of the protein could not be unambiguously assigned in the solid-state NMR-spectra.

In order to refine the characterization of the IXI-motifs, we carried out PDSD-measurements on wild-type αB and the valine mutants. To our initial surprise, the spectrum for the C-terminal mutant variant (**Figure 13**, red spectrum) showed signals at positions of peaks that were previously assigned by Jehle to I159 and I161. Since in this particular mutant I159 and I161 are not present, these signals could not arise from the C-terminal IXI-motif. The PDSD of the N-terminal mutant (**Figure 13**, green spectrum) resembles the wild-type spectrum (**Figure 13**, dark blue spectra at the bottom of the overlays) and as expected contains all signals previously assigned to the C-terminus. Only in the data of the quadruple mutant in which all four motif-related isoleucines are changed into valines all signals assigned to I159 and I161 are absent (**Figure 13**, light-blue spectrum). We therefore conclude that resonances of the N-terminal isoleucines I3 and I5 are indeed visible in the spectrum of the full-length construct but overlay with the signals deriving from I159 and I161. This clearly hampered their identification in the previous work by Jehle *et al.* Rather, the assignment from Jehle did not contain information about the three N-terminal isoleucines I3, I5 and I10, and it was speculated that they do not give rise to sharp cross-peaks as the N-terminus is known to be the most heterogeneous part of the protein.

From the current work, based on systematic mutations of both IXI motifs, we can now unambiguously conclude that the cross peaks which are observed in both the N- and C-terminal mutant, are at positions that originally were assigned to Ile 159 Cγ1-Cδ and Ile 161 Cγ1-Cδ as well as I159 Cγ1-Cγ2 and I159 Cβ-Cγ2. However, an attempt to assign the full stretch 2 – DIAIH – 6 using the original data from Jehle with the current knowledge of the assignment of I3 and I5 was not successful due to poor quality of the spectra and signal-crowding.
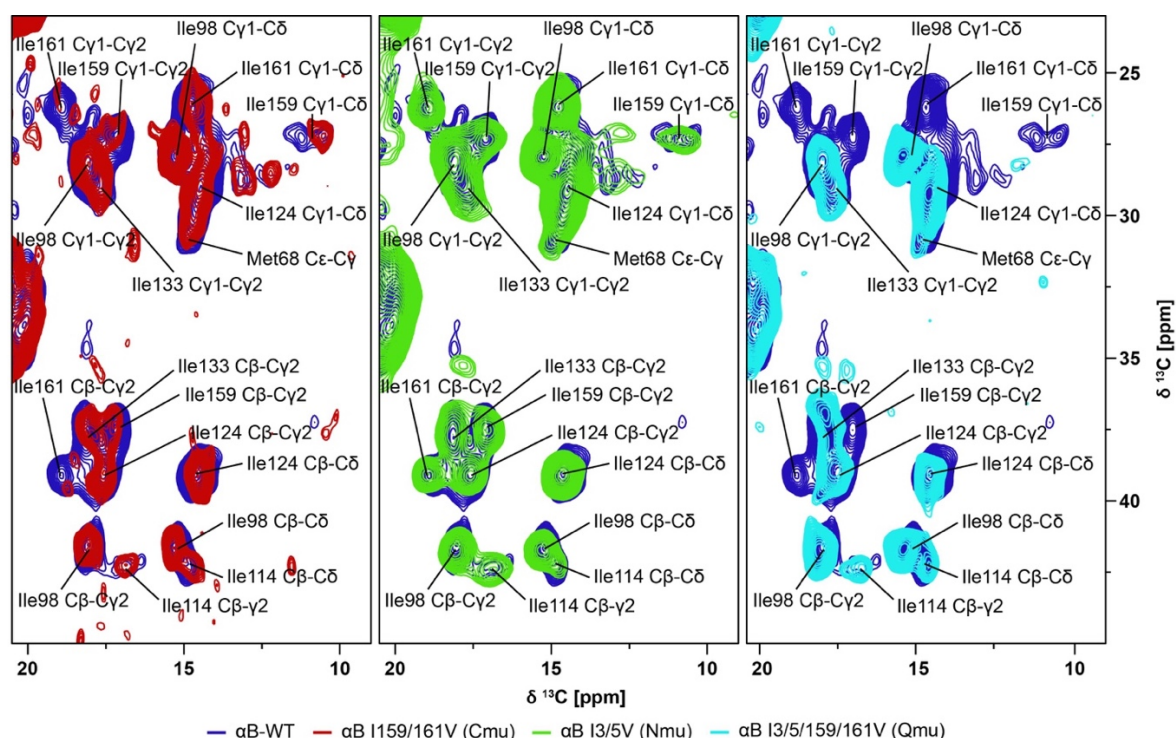
Figure 13: Overlay of the PDSD-spectra isoleucine-region recorded from valine mutants (Cmu red, Nmu green, Qmu lightblue). Bottom spectrum from αB-WT (darkblue). The assignment refers to the published chemical shifts by Jehle et al.[67].

The finding that the isoleucine side chains of both IXI motifs have very similar chemical shifts, strongly suggests that the side chains are in a very similar chemical environment. We know that for the C-terminal IXI motif the side chains are located in two hydrophobic pockets in the groove formed by β-strands 4 and 8. This unique environment results in a relatively large upfield shift of the side chain signals. Since multiple side chain signals of two isoleucines of the NTD have identical shifts as those of the CTD IXI motif, it is very tempting to conclude that the isoleucines of the N-terminal IXI motif reside in the same binding pocket. In other words, the fact that we have an overlay of signals to such an extend leads to the postulation of both IXI-motifs competing for binding to the same hydrophobic groove. The results from the PDSD-measurements on a set of mutants therefore confirm the assignment by Jehle for the C-terminal isoleucines and allow for the first time not only the assignment of the N-terminal IXI-motif but also provides a biological function for it.

One additional isoleucine, I10, could not be assigned by Jehle and this is also the case in this work. The spectrum of the quadruple mutant shows that I10 is not hidden

by signal-overlap with isoleucine residues from the motifs as the respective signals disappear completely in the light-blue spectrum.

To summarize, the Nmu-spectrum resembles that of the wild-type the most but there is a considerable amount of signal-intensity generated by I3 and I5 overlaying with signals assigned to I159 and I161. The N-terminal signals were not unambiguously assigned by Jehle due to the pronounced heterogeneity of the first part of the protein. The almost complete similarity in chemical shift-patterns for the isoleucines of both IXI motives strongly suggests that both IXI-motifs adapt a similar role in the oligomeric ensemble by binding to the same hydrophobic groove. In addition, it should be kept in mind that the isoleucines in both motifs are part of a different sub-sequence and therefore cannot take the exact same relative orientation; In the case of the N-terminal motif the "X" is an alanine, whereas in case of the C-terminus it is a proline. As the proline is the only cyclic amino-acid it forces its neighbors in either a cis or trans orientation towards each other. For every other spacing amino-acid the degree of freedom in orientation is much higher.

3.2.2 HSQC-experiments with αB-crystallin and valine-mutants

As Mainz *et al.* showed 2015[74], in HSQC experiments, exclusively the C-terminus is observed. The visible part ranges from residue S153 to the very end of the C-terminus (K175), including the IXI-motif. The reason that the CTD is observed is that it is highly polar and provides excellent solubility. Despite the CTD, the overall particle size of the average oligomer is fairly big and the overall complex too rigid and bulky to be fully seen by solution-NMR methods.

For the N-terminal mutant we obtain a HSQC-spectrum (light green spectrum in **Figure 14**) which is very similar to the wild-type spectrum. Compared to the originally published data from Mainz[74], the linewidth is considerably broader, which we attribute to the fact that the protein samples are not deuterated in our case. The overlay of the WT (dark blue) and Nmu (light green) HSQC-spectra reveal no distinguishable differences (**Figure 14**).

In addition, the spectrum of the Cmu (red) looks overall very similar to the WT (dark blue) as well, except for some expected new valine-signals and isoleucine-signals disappearing, accordingly. The effect of the mutation on neighboring residues is

comparably small, as can be deduced by the fact that the chemical shifts of T158 and T162 do not alter. A small difference in signal-intensity for certain amino acids like G154 can be recognized. This can be attributed to different experimental conditions. Interestingly, the spectrum for the Qmu (light blue) reveals more new signals than the C-terminal mutant. It is unclear if those signals arise either from residues which are generally not observed in the WT or if they belong to the second 'state' of the C-terminus as observed by Mainz[74], which could be populated to a higher degree in the quadruple mutant. This in turn would imply that in the Qmu the free C-terminus would be more populated than in the other two samples. The difference in the population could be very small and the change in the overall structure and oligomeric state only subtle. A third explanation for the signal missing in the WT-spectrum could be the general lower signal-to-noise ratio.
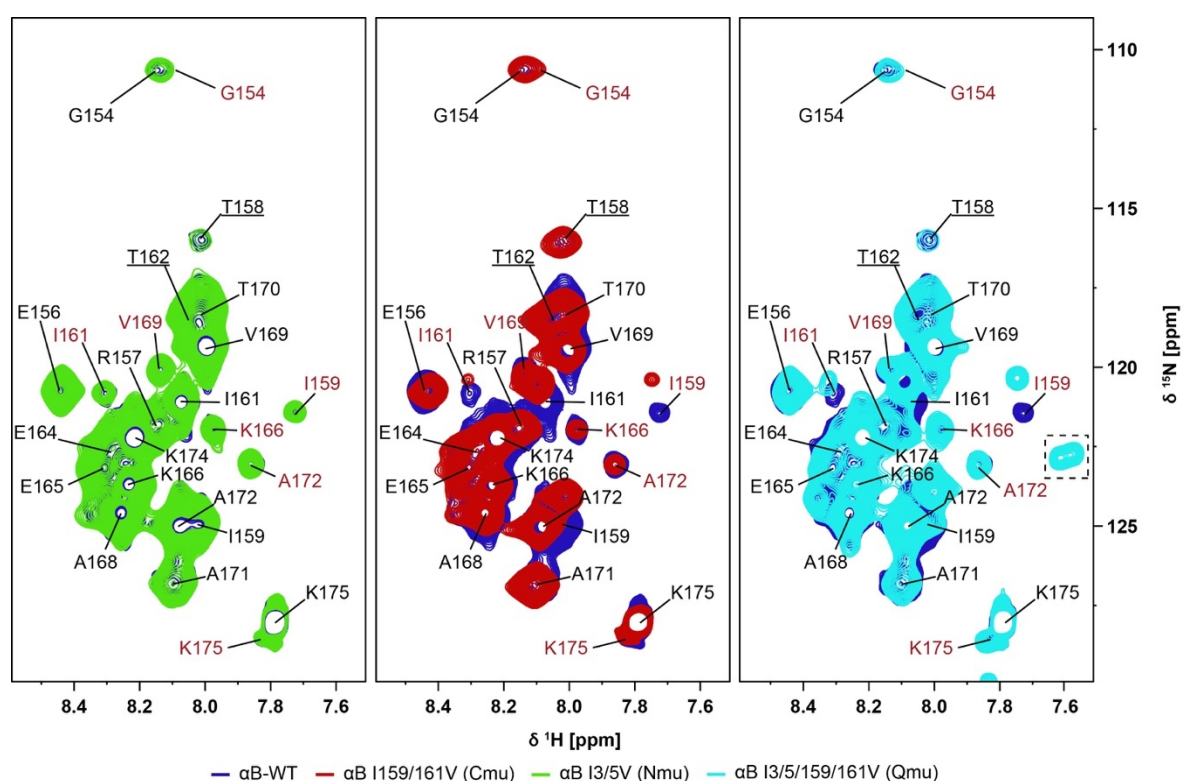
Figure 14: Overlay of the HSQC-spectra recorded from valine mutants (Cmu red, Nmu green, Qmu lightblue) plotted on top of αB-WT spectrum (darkblue). The assignment corresponds to the published chemical shifts by Mainz et al.[74]. Red labels indicate the second signal set.

Mainz showed that some residues have two signals with a significantly lower signal-intensity for the second set (assignment for second signal-set indicated in red in

Figure 14). In our case, due to the used labelling scheme and the resulting line broadening and overlap of peaks, less of those second signals can be unambiguously assigned compared to the published data of Mainz. Nevertheless, they are observed in the wild-type spectrum as well as in all mutant spectra and they derive most likely from the same oligomeric species as the major signals as it was shown through the similar translational diffusion coefficients[74].

3.2.3 ^{13}C-direct excitation in solution-NMR spectroscopy

The results from the PDSD- and HSQC-spectra display an apparent paradox. While in the solid-state the NMR-results suggest that the C-terminal IXI-motif exists in a rigid conformation, it is obviously highly flexible under conditions of solution NMR. The different sample conditions, in one case the sample being a solid precipitate and in the other case a dissolved protein, raises the question if the natural situation in terms of structure is properly displayed in the MAS-sample. To determine if the C-terminus is flexible at all in the precipitate, INEPT-based spectra were recorded under MAS-conditions. In a similar way Jehle has shown that the precipitated state still contains flexible parts in the ensemble[76]. The INEPT-spectra on the mutants look very similar to those reported by Jehle. The overlay (see Appendix **Figure 27**) of a CH-correlation obtained with solution NMR-spectroscopy with an INEPT-based CH-correlation from the MAS-sample displays a high degree of similarity in signals for the Cmu-sample. This leads to the assumption that the flexibility of the C-terminus in the solid-state sample is not lost through the sample preparation. Moreover, the observed binding of the C-terminal IXI-motif to the hydrophobic groove demonstrates that the motif can simultaneously occupy a bound and an unbound state. Furthermore, from the current assignments and the comparison between the different mutants, we can conclude that the isoleucine-signals observed in the CH-correlation and INEPT-based data derive from the C-terminal IXI-motif and not, as was assumed but not unambiguously assigned earlier[59,76], from the motif in the N-terminus.

Interestingly though, the INEPT-based CH-correlation of the C-terminal mutant contains several additional signals compared to the WT-, Nmu- or Qmu-spectra (Appendix Figure **Figure 28**). There is no obvious "spectroscopic" explanation for

this finding, so the reason might lie in subtle structural changes between the mutants. Although the mutations of isoleucines into valines do not result in a significant overall change in the oligomeric distribution, it could well be that the VXV-motif has a slightly decreased affinity to the hydrophobic groove. This could result in a subtle increase in the amount of flexible CTDs and therefore in a higher signal intensity for flexible C-terminal signals in the C-terminal mutant. Assuming that both termini compete in binding to the groove, this effect could lead to a subtle shift of the percentages of bound N- and C-termini between the two double mutants due to slightly different affinities, but would cancel out in the Quadruple mutant.

Signal integration could provide an estimate of the ratio between unbound and bound state of the C-terminal IXI-motif. However, the integration of signals in CP-based spectra and multidimensional solution-spectra is not a reliable method. Therefore, ^{13}C direct-excitation spectra without NOE built up were carried out on the solution-NMR samples. The resulting spectra (**Figure 15** for the WT) display the full protein, in which the rigid parts show up as broad 'background' and the flexible parts give rise to sharp signals.
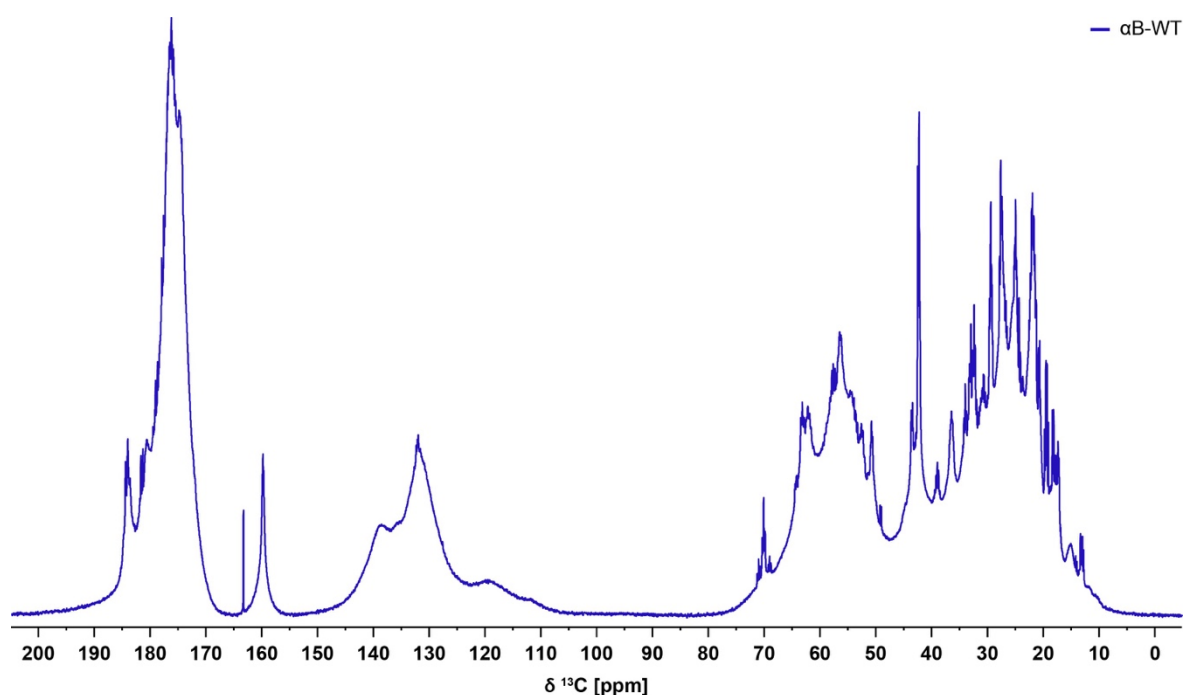
Figure 15: ^{13}C direct-excitation spectrum recorded in solution at 290 K from uniformly ^{13}C-labelled αB-WT. The delay-time before every scan was set to 30 s in order to allow for sufficient relaxation between the experiments.

The complementary properties of solid-state and solution-NMR in terms of type of detected signals is nicely reflected in the overlay of the 1D data with the PDSD (**Figure 16**).

In general, cross polarization-based spectra are well suited for investigations on rigid systems whereas the exact opposite is the case for solution-NMR experiments where only flexible domains or highly mobile proteins can be properly made visible. In the case of αB-crystallin, there are both, flexible and rigid regions in the oligomeric complex. Interestingly, the sharp signals at around 13 ppm in the 1D data overlay with the area of the PDSD where only small and broad resonances are observed. The other way around, areas of intense peaks in the solid-state data align with bulky areas in the solution-spectrum.
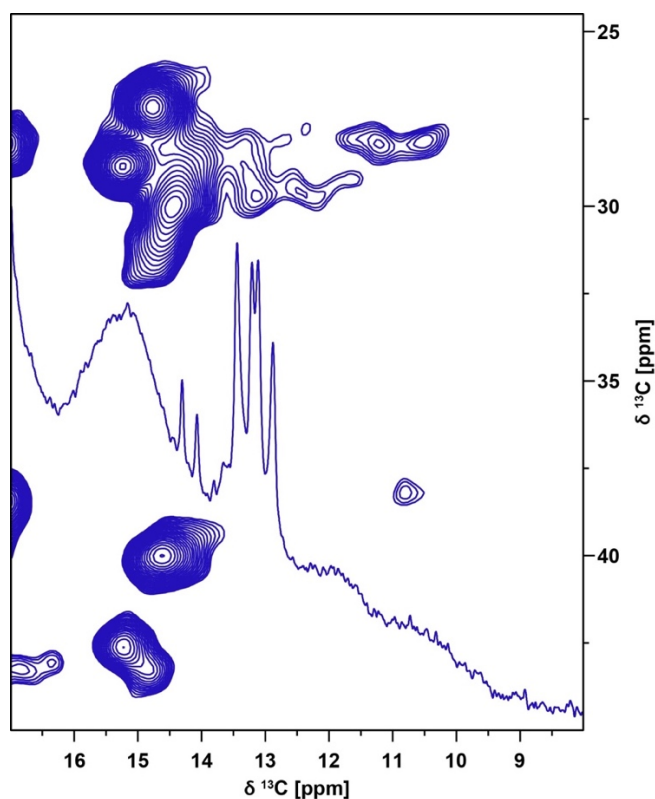
Figure 16: 2D PDSD MAS solid-state NMR-spectrum from precipitated ^{13}C-labelled αB-WT as "overlay" with 1D ^{13}C-direct excitation spectrum in solution, both recorded at 290 K.

From a comparison of the WT with the different valine-mutants it immediately becomes clear that there are several intense signals that disappear upon the mutation in the N-terminus. This is the case for signals all over the sidechain-area and in particular the chemical shift range where the isoleucine Cδs' are assigned (**Figure 17**; for full sidechain area, see Appendix **Figure 31**).

This clearly is in contradiction to the fact that the N-terminal residues do not give rise to cross peaks in the HSQC-spectra. From the currently available data it can only be hypothesized about the reason for this apparent paradox. Possibly, the sidechains of the respective residues are more flexible than the backbone. It could then well be that recording the HSQC and other multidimensional spectra with longer experiment times needed for unambiguous assignment, would reveal the cross peaks which could be assigned to the N-terminus. Alternatively, the mutation of I3 and I5 into valines could cause signal changes in the 1D data of residues in other parts of the protein which are structurally but not sequentially close by. However, in that case the

question why in the HSQC-spectrum of the Nmu and Qmu no signals disappear compared to the wild-type would still remain.
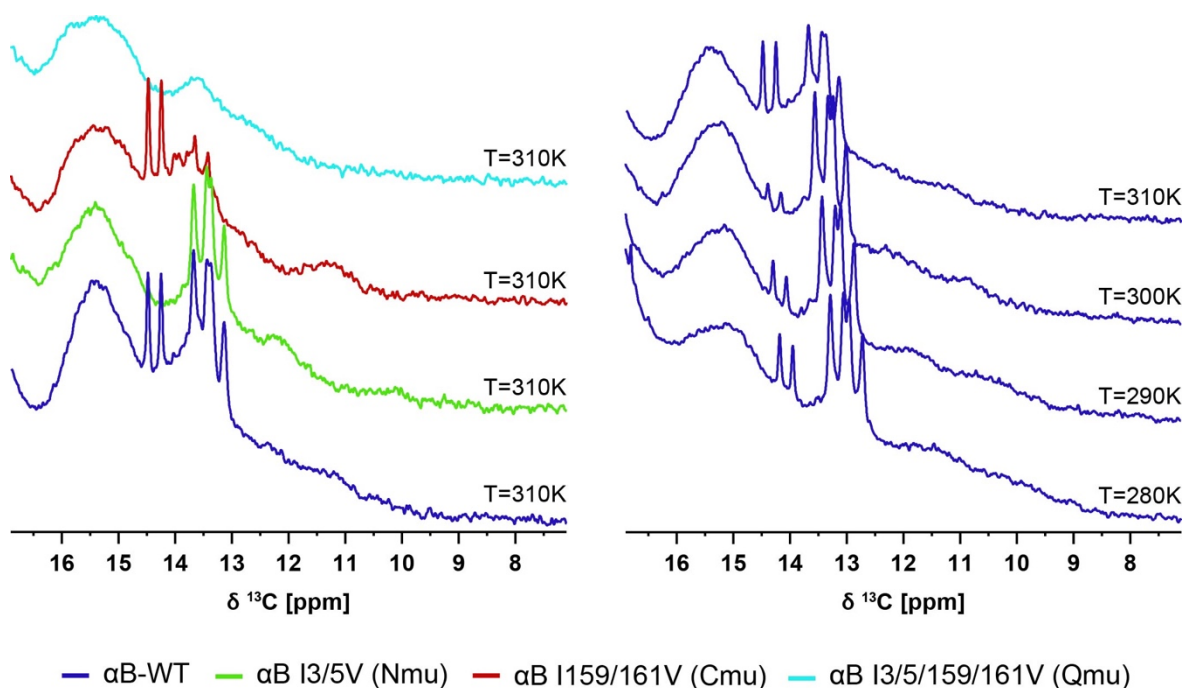
Figure 17: Left: ^{13}C-direct excitation spectra of ^{13}C-labelled αB-WT and valine mutants, recorded at 310 K. Color coding for the various mutants is indicated in the figure. **Right:** ^{13}C-direct excitation spectra of ^{13}C-labelled αB-WT recorded at 280 K, 290 K, 300 K and 310 K. The delay-time before every scan was set to 30 s in order to allow for sufficient relaxation between the experiments.

In an attempt to achieve more resolution in the low chemical shift range, a 1D direct-excitation spectrum of a 2-^{13}C-glycerol labelled WT-sample was acquired (Appendix **Figure 29**). For this, a sample was prepared by use of 2-^{13}C-glycerol as only carbon source in the recombinant protein production. In the resulting sparse labeling, the Cδ of the isoleucines is ^{13}C labeled and connected to an unlabeled Cγ (for labelling scheme see Materials & Methods, **Figure 10**). Hence, this removes the *J*-coupling between the Cγ and Cδ of the isoleucines. In the upfield chemical shift area of the one-dimensional data, this results in a change from doublets into singlets with higher intensity. Interestingly, one of the doublets which was assigned to an N-terminal isoleucine does not change into a singlet but disappears (marked with * in **Figure 18**). This in turn leads to the conclusion that the corresponding doublet in the WT- and Cmu-spectrum cannot derive from an isoleucine but from another type of

residue. It is not unlikely that it comes from an alanine Cβ, which would be in line with the labelling pattern for alanines for the 2-^{13}C-glycerol type of samples.
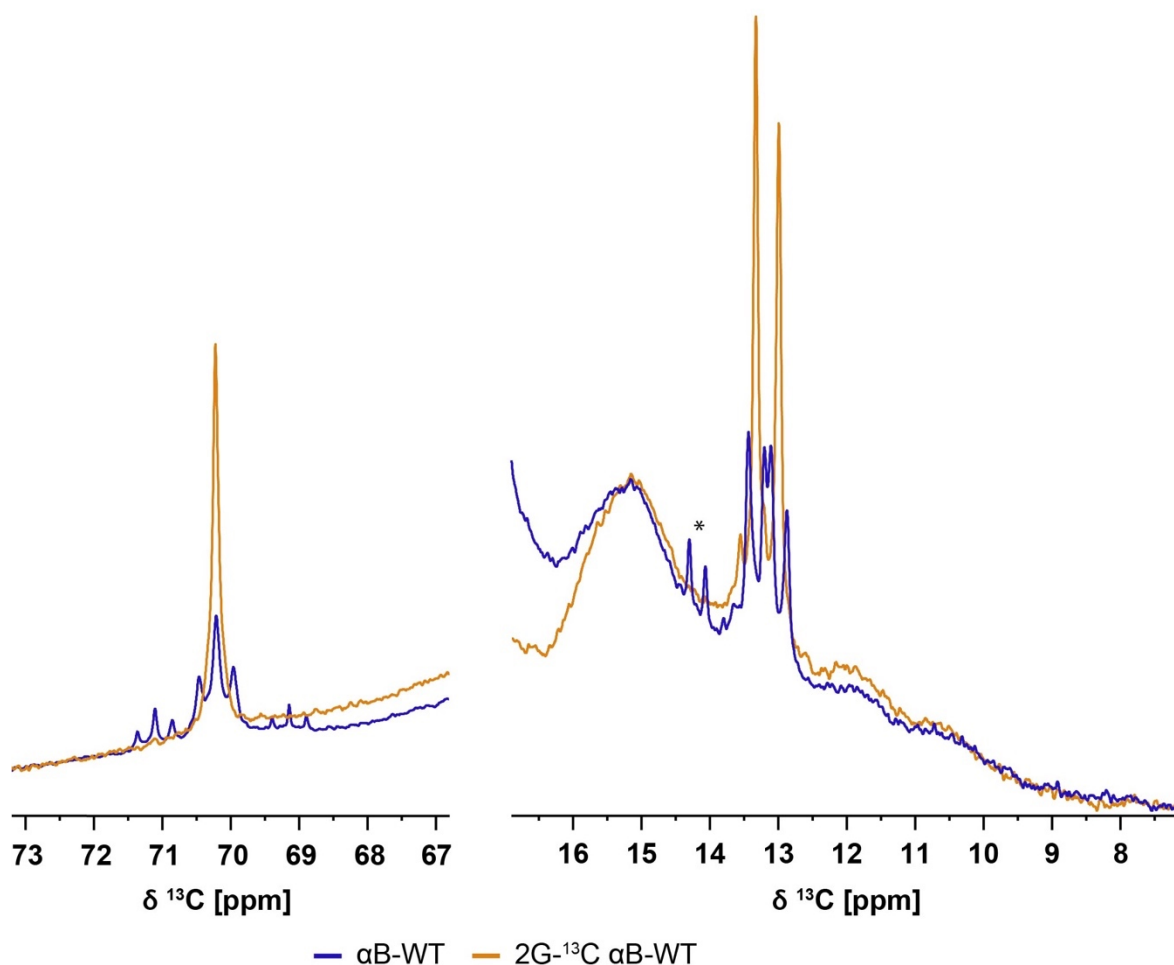
Figure 18: Selected regions of ^{13}C-direct excitation 1D spectra of uniformly 13C-labelled αB-WT (dark blue) and 2G-^{13}C-labelled αB-WT (orange) at 290 K. The delay-time before every scan was set to 30 s in order to allow for sufficient relaxation between the experiments.

In the HC-INEPT data of the mutants (**Figure 28**) a methionine Cε-signal could be identified. As αB possesses only two methionines, in positions 1 and 68, and the latter one is situated at the beginning of the rigid core domain, it is deduced that this signal derives from M1. Although the beginning of the N-terminus could not be assigned in HSQC spectra from αB-WT or the valine-mutants, due to the observed resonances for Cε from M1 and the fact that the ^{13}C-direct excitation data contains a sharp alanine-signal, we tentatively assigned this sharp signal to A4, the spacer in the N-terminal IXI-motif.

The uniformly labelled wild-type spectrum contains three triplets at around 70 ppm, with the triplet in the middle having a three-fold intensity compared to the outer ones (**Figure 18**). Usually threonines are found in this particular ppm-range. As there are three such triplets one could conclude to assign those to the three threonines, T158, T162 and T170. As it turns out, the outer triplets disappear in the sample of the sparse labelling which should not be the case according to the labelling scheme. An assignment to the Cα of histidines H6 and H7 would be in line with the idea that sidechain-atoms in the very beginning of the N-terminus are flexible enough to give rise to sharp peaks in the direct excitation spectra. A chemical shift of around 70 ppm is unusual but not impossible for histidine α-carbons, according to the chemical shift statistics in the BMRB.

In order to further assign the peaks in the 1D-data for all four mutants, CC-TOCSY spectra were recorded, as well as several other multidimensional spectra for the WT and the Cmu (For exact list see Materials & Methods section 2.2.3). The assignment of resonances is comparably difficult due to the lack of dispersion of signals which derives from the unfolded character of the flexible CTD. The parts that could be unambiguously assigned are in agreement with the published assignment by Mainz[74], except for changes related to the mutations (for the chemical shift list see Appendix, **Table 11** to **Table 13**). The strongest triplet at 70 ppm in the ¹³C direct-excitation data indeed comprises the α-carbon signals of all three C-terminal threonines, hence for these particular atoms the extend of overlay is basically 100%. However, the smaller signals that we tentatively assigned to H6 and H7, do not show cross-peaks in the CC-TOCSY spectra of the WT or Cmu. Beyond that, none of the sharp signals in the ¹³C direct-excitation spectra of WT and Cmu which are attributed to the N-terminus show cross-peaks in the CC-TOCSY spectra.

The different behavior with respect to signal intensities for certain signals is observed for the different mutants. For instance, the two C-terminal IXI-isoleucines I159 and I161, can be identified in the WT spectrum and their signals are significantly less intense in the Nmu-spectrum. Again, a possible explanation could lie in a subtle change in affinity for different forms of the V/I – X – I/V motif. If the IXI motif binds slightly stronger than a VXV motif, the Nmu would have slightly less free C-terminal motif which would result in a decrease in signal intensity.

To assess a temperature dependent behavior of the individual signals, ^{13}C-direct excitation experiments were carried out at different temperatures (**Figure 17**). Comparison of the resulting spectra revealed that the signals behave differently with the temperature. For some sidechain signals, no significant changes in relative signal-intensity as function of temperature can be observed. However, another set of signals experience a minimum in intensity at 300 K. This includes the A4-doublet at 14 ppm and the Cα-signals of H6 and H7. It is worthwhile to mention, that all signals which disappear in the spectra of the N-terminal mutant show this particular temperature dependent behavior. In contrast, other signals, like the I159/I161 Cδ-doublets exhibit a complementary temperature dependency and show a maximum in intensity at 300 K.

Due to the absent baseline resolution of signals and the large number of isoleucines in αB-crystallin (nine out of 175 amino acids are isoleucines), a direct integration of signals is not possible. Nevertheless, if the 1D spectra are subtracted from each other, the difference spectrum displays only the signals deriving from the particular isoleucines and valines which were exchanged upon mutation. This approach towards integration of signals is valid as long as the structural changes upon mutation are negligibly small so that the bulk or background contribution from all other residues to the signal will be highly similar. In that case, the difference between the WT and the N-terminal mutant will consist of all signals deriving from I3, I5, V3 and V5, with the isoleucines appearing as positive signals and the valines as negative. For the difference between WT and Cmu we have the respective situation for the residues 159 and 161, in the difference between the WT and Qmu for all four of them. To confirm that no significant structural changes took place upon mutation, the differences between the Qmu- and Cmu- and accordingly, Qmu and Nmu-spectrum should overall look highly similar. The expected signals with the attribution to the specific residues is summarized in **Table 7.** The spectra which should correspond to each other are color-coded in the same way, and reflects the color coding used for the different mutants so far: Red is chosen for the visible isoleucines of the N-terminus, green for the isoleucines of the C-terminus and dark blue where all motif-related isoleucines are observed like in the WT.

Table 7: Overview of subtracted ^{13}C direct-excitation spectra with expected residing signals, for different combinations of WT and mutant αB spectra.

Subtracted spectra	Expected **positive** signals	Expected **negative** signals
ΔWT/Nmu	I3 and I5	V3 and V5
ΔWT/Cmu	I159 and I161	V159 and V161
ΔWT/Qmu	I3, I5, I159 and I161	V3, V5, V159 and V161
ΔCmu/Qmu	I3 and I5	V3 and V5
ΔNmu/Qmu	I159 and I161	V159 and V161

In **Figure 19** the resulting difference-spectra for 1D experiments of the various mutants recorded at 310K are displayed. The pair-wise highly similar spectral shapes of the corresponding subtractions, implies that the structural changes through the mutations can be neglected.

In the subtracted spectra, baseline resolution is not fully accomplished for all signals, but nevertheless integration of distinct signals is feasible. For the correct interpretation of the integration values, it is essential to recapitulate the assignment: From the comparison of the valine-mutant spectra with the WT and characterization through the 2-^{13}C-glycerol spectrum, the doublet at 14.4 ppm in the red and dark blue spectra is attributed to an alanine-signal and assigned to A4. The two sharp doublets at 13.2 and 13.5 ppm in the light green and dark blue spectra derive from the I159 and I161 Cδ. The assignment is indicated in **Figure 19**.

Interestingly, in the same ppm region of these two doublets in the red spectrum, there are signals which are not as sharp as the ones from the alanine but do not appear as completely broad. Those signals are hypothesized to derive from the N-terminal isoleucines I3 and I5. As all those signals appear as sharp they are attributed to the unbound portions of the IXI-motifs. The comparison of the WT in the 1D with the PDSD (**Figure 16**) implies that the broad peaks to the right end of the difference-spectra represent each one isoleucine of the two IXI-motifs in the bound form. For the C-terminal motif this is assigned to the I159[68], for the N-terminal motif an assignment to a specific residue in the motif, either I3 or I5, is not unambiguously possible with the given data.
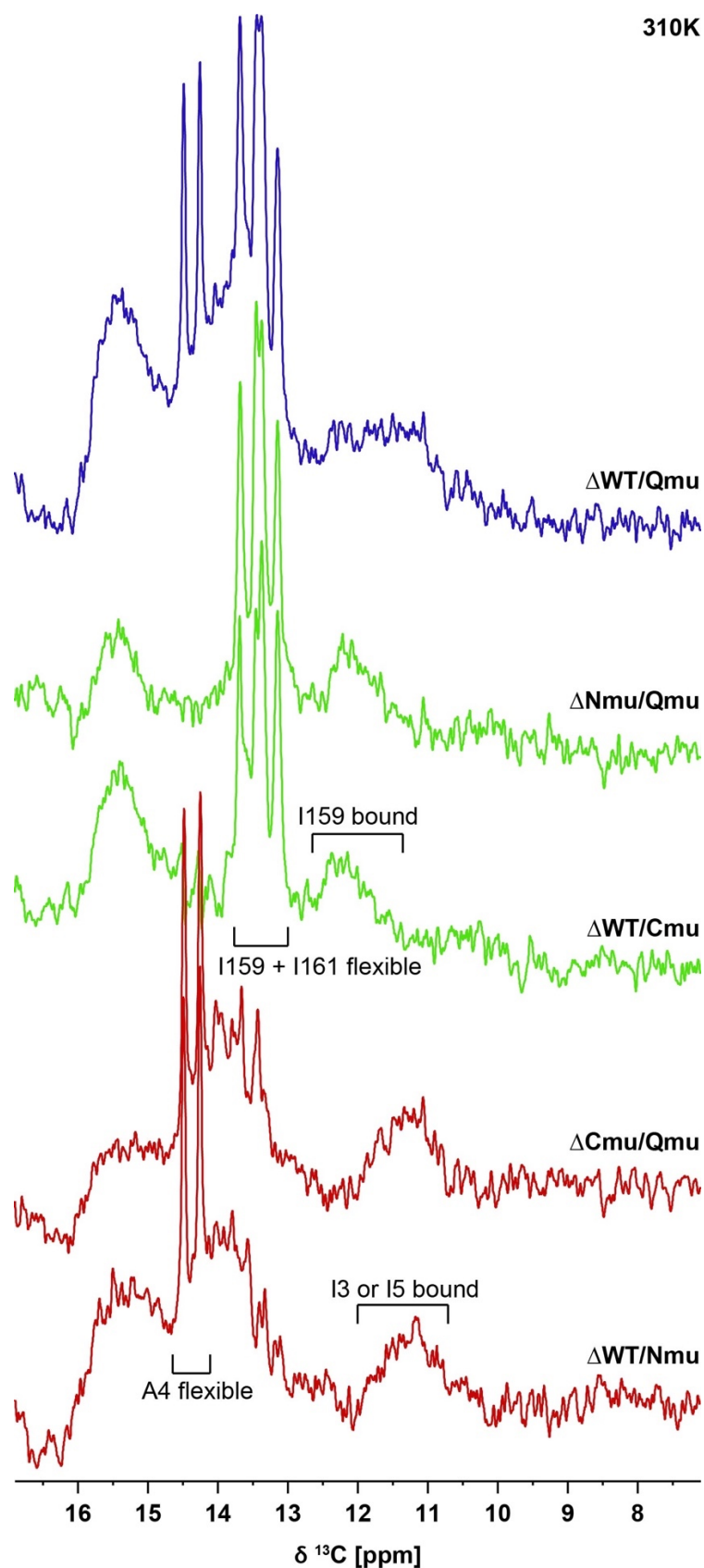
Figure 19: Difference-spectra obtained by subtraction of respective ^{13}C direct-excitation spectra recorded of the various mutants at 310 K. Color coding refers to the remaining motif-related positive isoleucine-signals resulting in dark blue for ΔWT/Qmu as all four motif-related isoleucines give positive signals, light green for positive C-terminal isoleucines, red for positive N-terminal residues. Assignment indicated for selected discussed signals.

The integrals for the different peaks in the spectra obtained from subtraction from the WT at all measured temperatures (290 K, 300 K and 310 K) are given in **Table 8.** The subtracted spectra at 290 K and 300 K and the regions used for the integration are shown in the Appendix, **Figure 32**, **Figure 33** and **Table 9**.

The values are normalized to 1.0 for the broad peaks deriving from the immobilized species on the right end of the spectra as those peaks are best baseline resolved and only one isoleucine contributes to them in each case, I159 and I3 or I5, respectively. The assignment is discussed above. The values obtained from the subtraction of the WT with the Qmu serve as a control, as they display the sum of ΔWT/Cmu and ΔWT/Nmu.

The integration values convincingly show that the ratio between unbound and bound termini behaves complementary for the different termini. The ratio of unbound to bound C-terminus can be derived from the values obtained for the difference spectra of WT and Cmu (green in **Table 8**). It is around 2 to 1 at 290K, keeping in mind that two isoleucines (I159 and I161) contribute to the sharp signals which cannot be separated for the integration, but only I159 gives rise to the broad signal. With rising temperatures, the ratio of unbound to bound IPI-motif changes to ~ 1/1 at 310 K. For the N-terminal motif the temperature-dependency of the bound/unbound-proportions changes exactly vice versa as deduced from the ratios of unbound alanine-signal (A4) to bound isoleucine (I3 or I5). It goes from 0,7/1 at 290 K over 0,9/1 at 300 K to 1,7/1 at physiological temperature.

Notably, the values do not exhibit the amount of possibly uncovered hydrophobic groove and do not claim to represent all possible forms of the termini, especially not for the N-terminus.

Table 8: Integration values obtained from Difference-spectra (Figure 19, Figure 32 and Figure 33). Exact integration areas in ppm are listed in Table 9.

subtracted spectra	T[K]	unbound Alanine (A4)	unbound Isoleucines (I3/5/159/161)	bound C-terminal Isoleucine (I159)	bound N-terminal Isoleucine (I3 or I5)
ΔWT/Nmu		0,69	1,31	-	1,00
ΔWT/Cmu	290	-	2,90	1,00	-
ΔWT/Qmu		0,70	4,22	2,00	
ΔWT/Nmu		0,91	1,96	-	1,00
ΔWT/Cmu	300	-	2,08	1,00	-
ΔWT/Qmu		0,85	4,17	2,00	
ΔWT/Nmu		1,68	2,44	-	1,00
ΔWT/Cmu	310	-	1,94	1,00	-
ΔWT/Qmu		1,57	4,45	2,00	

In the attempt to recombinantly produce αB-crystallin, Mainz tried to purify it with use of a His-tag[74] and proteolytic cleavage of the tag with enterokinase. This resulted not only in pure protein without the His-tag, but as well in a certain proteolytically digested shorter construct lacking the C-terminus and both variants still with His-tag on it, even though a reaction time of 40 h was applied. It turned out that the enterokinase, although in general a very specific kinase, unspecifically cleaves the C-terminus after residue R157. The cleavage of the C-terminus took place to the considerable amount of approximately 50% according to the interpretation of Mainz. In our case we can reproduce this result with all our mutants (**Figure 34**). Although the subunit-exchange in the oligomeric ensemble takes place on a minute time-scale[74,152], approximately 50% of the C-termini seem to be safe from proteolytic cleavage with enterokinase. Our results from the integration of the subtracted spectra is perfectly in line with the cleavage result assuming that only unbound C-termini are accessible for the proteolysis.

3.3 Site-directed spin-labelling of αB-crystallin

In protein-NMR site-directed spin-labelling can be applied to obtain large scale, global information about distances in proteins[153]. Through the introduction of a paramagnetic spin-label into a protein by covalently attaching it to a cysteine, NMR-measurements can make use of the paramagnetic relaxation enhancement-effect (PRE). The magnetic moment of the paramagnetic center, in the current study these are unpaired electrons, is several orders of magnitude larger than the magnetic moments of the nuclei of the atoms. As the paramagnetic center only couples strongly with nuclei in spatial proximity, this results in enhanced relaxation in a distance-dependent manner[154–157]. This effect reaches out to long-range distances up to 20 Å, whereas NMR-couplings typically allow direct interactions over distances of only a few Ångström[153].

3.3.1 Spin-labelling reaction

In order to substantiate our results that the N-terminal IXI-motif can also bind to the hydrophobic groove, we aimed to spin-label selected cysteine-mutants of αB-crystallin. The wild-type protein does not contain any cysteines that can be used to attach a spin label at a defined position in the protein. To this end, we used the mutant-variants αB A4C and αB P160C where the spacer-residues X in the IXI-motifs is mutated to a cysteine in order to be able to attach a paramagnetic label in a chemical reaction as shown in **Figure 20**. As the sidechain of the spacer is oriented in the opposite direction as that of the isoleucine-sidechains, it is expected that the mutation does not alter the binding of the motif in a significant way. This particular positioning of the spin-label is chosen because of the direct proximity to the relevant isoleucines and because it allows to investigate the termini independently. It was prescinded from the introduction of a spin-label in or close by the binding groove itself, because the effect on binding could not be foreseen and the independent investigation of the termini would not be possible.

As first step for the spin-labelling reaction, the buffer of the mutant protein-solution (αB A4C or αB P160C) needed to be exchanged into a dithiothreitol-free buffer. As a reducing agent DTT is responsible for keeping the sulfhydryl-group of the cysteines

in the reduced form in order to avoid disulfide-bond formation. If not removed from the solution it would prohibit the reaction of the spin-label with the protein or cleave it off immediately after bond-formation.
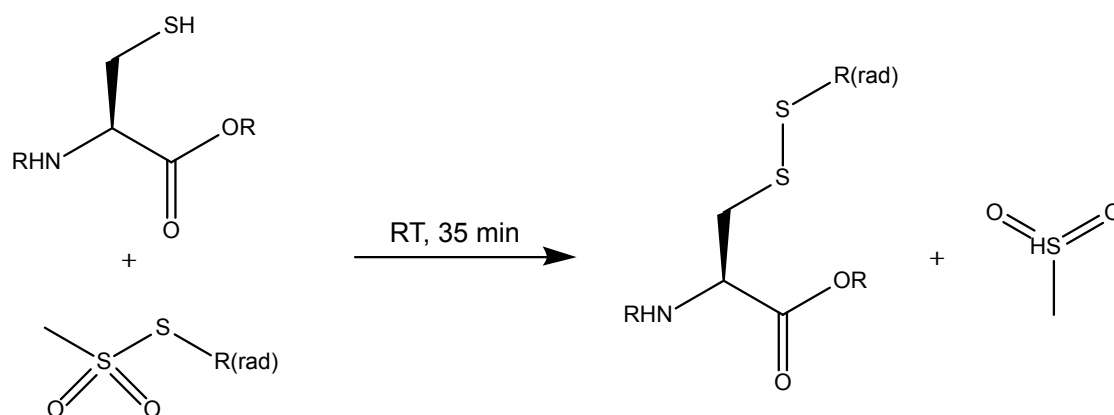
Figure 20: Chemical equation for incorporation of the paramagnetic spin-label via reaction with the sulfur group of a cysteine.

3.3.2 Spin-labels for αB-crystallin

The strength of the PRE-effect is a function of the distance from the paramagnetic center, in this case the unpaired electron. Typically, MTSL (**Figure 21**) is used for this purpose. From previous NMR studies it is known that the a complete vanishing of signals will take place for distances up to ~5 Å from the label; in the range of 5-15 Å the signal reduction is incomplete, and beyond ~15 Å no relaxation enhancement should be observed[153]. To further extend the radius of the PRE-effect it is possible to use a longer linker. However, this is accompanied by more possible orientations of the radical to the protein and which makes interpretation of the resulting NMR-data more difficult. Another option to extend the effective range of the PRE-effect in the protein is the use of a biradical. To this end we used a TOTAPOL-derivative (**Figure 21**), a radical commonly used in dynamic nuclear polarization applications[158–160]. The bleaching radii of the two radical-centers should overlay. In this work we used both type of radicals, MTSL and the TOTAPOL-derivative, for the investigations on the terminal IXI-motifs of αB-crystallin.

MTSL

TOTAPOL-derivative SB-II-18
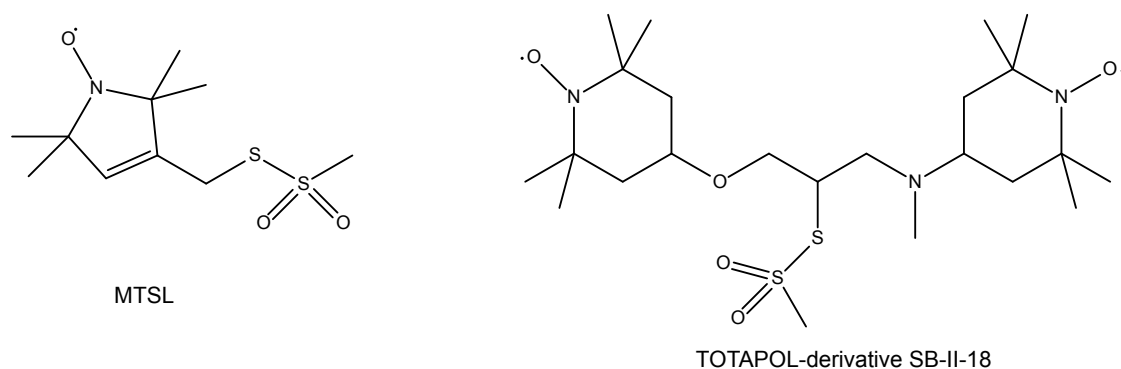
Figure 21: Chemical structures of MTSL and TOTAPOL-derivative SB-II-18 used for spin-labelling of αB-crystallin cysteine mutants αB P160C and αB A4C.

3.3.3 EPR-spectroscopy with spin-labelled αB-crystallin

Electron paramagnetic relaxation spectroscopy measurements were carried out for characterization of the spin-labelled samples. The labelling efficiency as determined by EPR was different for the two mutant variants. For the αB P160C-MR (monoradical) variant the efficiency was 75%, for the αB A4C-MR it was only 34%. Moreover, the EPR measurements revealed that both samples lose their label at room temperature. Whereas for the P160C-variant, free radical starts to get visible after more than a day of storage at RT, for the A4C sample this happens much faster and signals for free radical grow on a timescale of only a few minutes.

The polydisperse and heterogenous character of the system was clearly reflected in the DEER results, both for the N- and C-terminally labelled proteins. DEER measurements were performed to determine the distances between the radicals in the system. They revealed a very heterogenous distribution of radicals (see Appendix **Figure 35**). The DEER curves provide an explanation for the observation in the EPR-results that the N-terminally labelled protein quickly loses its label. According to the DEER results, there is a considerable portion of label closer to each other than 2 nm. Nevertheless, the DEER data should be interpreted with some caution because of the poor quality of the curves and the apparent variabilities between the DEER-traces obtained with different acquisition times, which is a sign of poor reproducibility.

Unfortunately, the quick loss of the paramagnetic label on a minute timescale in the αB A4C sample prohibits any sensible solid-state NMR applications as the sample

preparation and the experimental time per spectrum already accounts for more than two days. Under these circumstances, no reasonable amount of spin-label would remain attached to the cysteine in the protein in the course of the measurement.

The quick loss of the label when attached to the N-terminal IXI motif can be explained by both models for the 24mer; According to the Weinkauf-model, the N-termini are mostly involved in formation of the 3-fold symmetry bridge-heads which connects the barrel-like core domain structures. Half of the N-terminal IXI-motifs come close to each other with distances < 3 nm. For these, the cleavage of the label and formation of disulfide-bonds between the cysteines increase to the entropy of the system, which provides an excellent unidirectional driving force for the reaction. In the Jehle-model, the situation is even more evident as all N-termini reside towards the inner cavity and the cysteine residues are in close proximity to each other, with maximum separation of 1.2 nm to at least two other mutation-sites. In this model, the low labelling-efficiency becomes reasonable, as the cysteines can readily form disulfide-bonds with each other, instead of reacting with the radical which first has to reach the inner cavity of the chaperone. In either case, both models give a plausible explanation for the poor labelling efficiency and the quick loss of the radical for the N-terminal IXI motif.

3.3.4 NMR-spectroscopy with spin-labelled αB P160C

3.3.4.1 HSQC-experiment with spin-labelled αB P160C

Figure 22 shows HSQC spectra of spin-labelled and unlabeled αB P160C (black and red assignments refer to the WT major and minor signal set, green assignments to αB P160C), overlaid with spectra from WT αB-crystallin. The linewidth in the HSQC spectrum with spin-label is comparably broad but the bleaching effect is concentrated at the area around the labelled residue. Since the flexible portion of the C-terminus visible in HSQC-spectra lacks a tertiary structure the clustering around residue 160 is in line with the expectations. Additionally, no shifts for visible residues are observed.
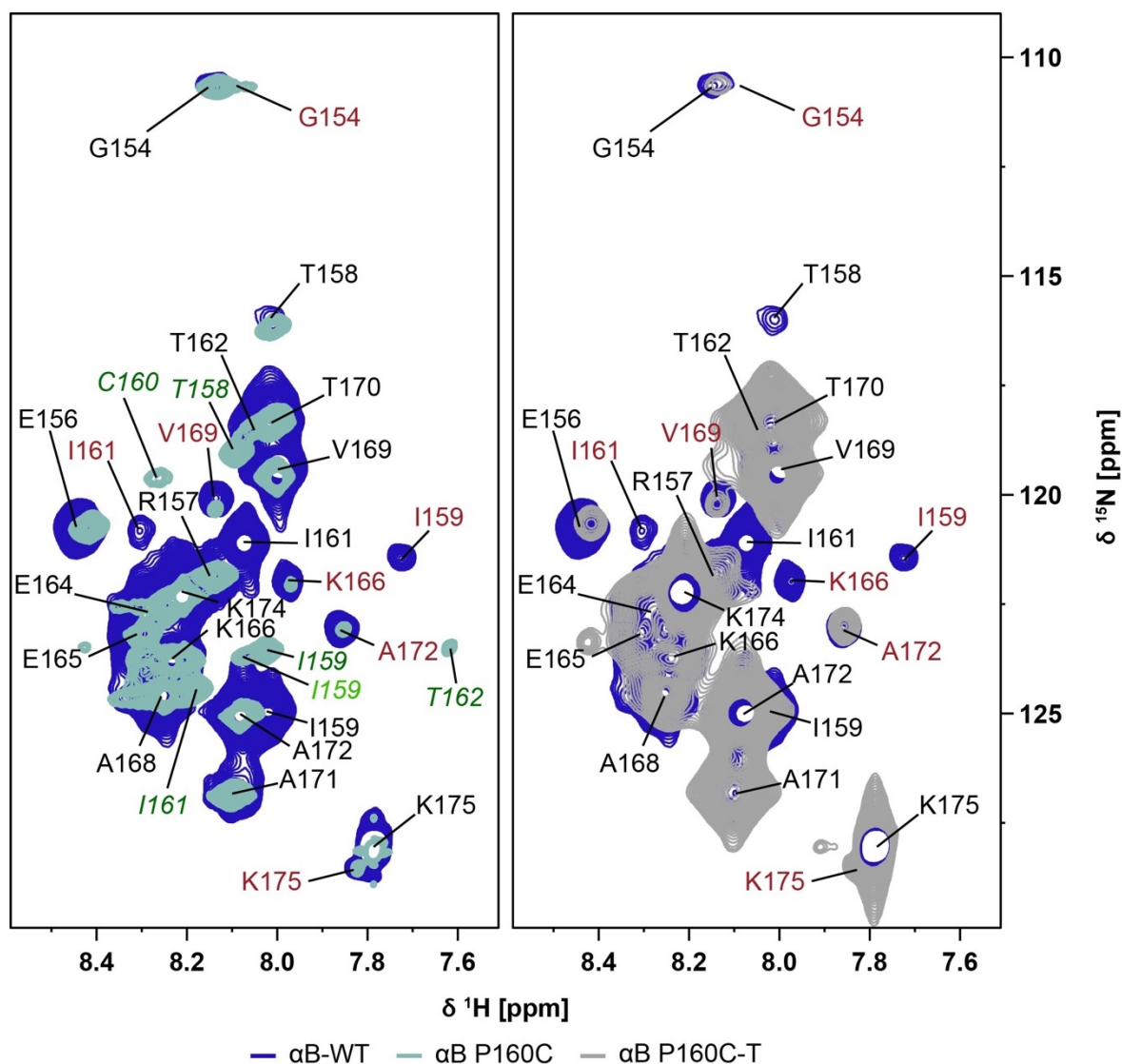
Figure 22: Overlay of HSQC spectra of unlabeled αB P160C (turquoise) and TOTAPOL-labelled αB P160C (light grey) on top of αB-WT (blue). The assignment in black for the major set of signals and red for the minor set of signals, refers to wild-type spectrum and corresponds to the published chemical shifts by Mainz *et al.*[74]. The assignment in dark green italic refers to the major signal-set of αB P160C carried out within this work, the light green italic assignment is attributed to the minor signal-set in αB P160C. For experimental parameters see Table 4 in section 2.3.1.

3.3.4.2 PDSD-experiment with spin-labelled αB P160C

The evaluation of the PRE-effect was done via signal comparison in PDSD spectra acquired with similar parameters (temperature, mixing time, number of scans) as used for the wild-type spectrum. In general, the overall signal intensity decreased compared to the wild-type spectrum. As expected, this was the case to a higher extend for the bi-radical than for the mono-radical (Appendix Figure **Figure 36**).

The chemical-shift list for αB-crystallin from Jehle (*BMRB* Entry 16391) served as basis for the classification of spins into four categories: The first category is *red* – this

category comprises every spin for which the signal experiences strong until complete vanishing. In general, 'complete' is a bit of an unfortunate criterion as the labelling-efficiency was not 100% and we assume that not every C-terminal IXI-motif with spin-label is embedded similarly in the hydrophobic groove within the oligomers. Nevertheless, the general decrease in signal-intensity can still contribute to complete vanishing of signals. On the other hand, line broadening of signals that are already weak could render them undetectable, even though these spins are at a relatively large distance from the spin label. The second catagory is *orange* – which includes all spins for which the signals experience a recognizable PRE-effect but are still visible in the spectra. The third category, *green*, includes all spins for which the signals are not affected by the PRE-effect. The overall decrease in signal-intensity makes the classification as green difficult; in general, "*green*" denotes those signals that appear measurably stronger compared to the affected ones. All residues not falling into one of these first three categories, for instance because they cannot be resolved due to overlay with other signals or because of lacking assignment, are categorized as *grey*. Distances were determined by mutation of residue 160 to an amino acid with a long side chain (in our case we chose arginine) within *PyMol* and measured from the amid nitrogens to the closest atom of the affected residue. The spacer length equates to the length of the chain from Cα of residue 160 to the tip of the radical.

3.3.4.3 Paramagnetic relaxation enhancement (PRE) – effect of mono-radical vs. bi-radical discussed at 24mer models

For comparison of the PRE-affected spins between the two radicals, the result of the evaluation of the PDSD spectra is mapped on the two 24mer-models of αB-crystallin (**Figure 23** for effect mapped on the Weinkauf-model, **Figure 24** for mapping on the Jehle-model).

As the spin-labelled proteins are measured as solid precipitates, it is likely that the PRE-effect of a radical attached to a monomer within the oligomer, reaches out to neighboring oligomers. However, we assume that this effect happens with a statistical distribution over all atoms close enough to the oligomer surfaces; hence,

we expect it averages out to an overall line broadening effect but can be neglected in the discussion of the PRE-effect within an oligomer.

A general issue is the relatively high number of residues that is categorized as *grey*. The assignments for N-terminal residues are very sparse and in general the PDSD spectra lack the resolution for unambiguous identification of many amino-acid types. As a result, only residues that have a characteristic chemical shift distribution, such as threonines, alanines, glycines, prolines and isoleucines, can be unambiguously identified whilst residues without characteristic chemical shift patterns remained unassigned and classified as "*grey*".

The *green* residues form mainly two clusters in both models. One in the form of a stripe across the α-crystallin domain, the β-barrel-like arrangement of β6+7 and β5. The second cluster comprises N-terminal residues, which are in case of the Jehle-model surface-exposed. In the Weinkauf-model the second cluster is embedded in the interior of the bridge-heads formed by the N-termini.

In the Weinkauf-model most of the residues belonging to the *orange* category cluster systematically around the mutation-site, with G64 and A57 being the furthest away by 14 and 17 Å, respectively. Nevertheless, the close distance to A57 is only for the inner C-termini, because for the IXI-motifs located at the outer edge of the barrel-like core the distance is fundamentally longer (>34 Å). In the Jehle-model, the orange-categorized residues are localized mostly close to the residue 160 or in the area of the loop-interface between two core-domain dimers. This area is in a sandwich-position between two C-terminal IXI-motifs, with one located at the outer edge of a core domain beta barrel and the other at the inner end of the respective adjacent crystallin-domain dimer. Only A57 and G4 form an exception, being relatively far away from the mutation site and accordingly, to the radical.

For the mono-radical, not many signals can be categorized as *red*. The only residues in this category are the isoleucines directly next to the labelling-site and the three residues P86, A126 and I133. P86 is localized in the loop between β3 and β4, A126

and I133 in the loop between β6+7 and β8. This means that all those residues are situated in unstructured areas of the protein, a fact which is accompanied with a certain mobility. In the Weinkauf-model, P86 is in direct proximity to the C-terminal IXI-motif. A126 is with 12-15 Å relatively far away, but still within a reasonable distance from the radical to experience a PRE-effect. In the Jehle model, P86 is also relatively close to the radical. On the other hand, the distance between the radical and A126 is approximately 20 Å and hence quite large to explain such a strong PRE-effect. Nevertheless, a judgement solely based on the measured distances from the amide nitrogen of an arginine residue in position 160 to the closest atom of the respective amino acid can only serve as a rough estimation, since the exact positioning of the spin-label is not known due to spatial flexibility of the linker chain.

For the bi-radical, the PRE-effect is more pronounced as evidenced by the fact that almost three times as many residues are classified as *red* compared to the mono-radical. Many of the residues which have been marked as *orange* in the mono-radical evaluation changed to the *red* category (all together 9 residues), which is in line with the expected increased PRE-effect due to radical structure and geometry. A57 moved from *orange* to *red* but, as discussed above, is located relatively far away from the radical in the Jehle-model. In general, the red residues in the bi-radical sample cluster systematically around the IXI-motif in the Weinkauf-model.

Remarkably, the stretch of *green* residues in the α-crystallin domain is only partially changed into *orange*. Residues involved here that change to *orange* are E117 in β6+7 and G102 in β5, but not K103 and H104 in the same β-strand.
The *green* cluster in the N-terminal region essentially remains similar as before, more surface oriented in the Jehle-model and embedded in the center of the bridge-heads formed by the N-termini for the Weinkauf-model. The outreach of bleaching-effect in the biradical sample is reasonable as the linker from the disulfide bond to the radical is longer than in MTSL (eight vs. five bonds, respectively). Most likely, steric effects play a role as well.
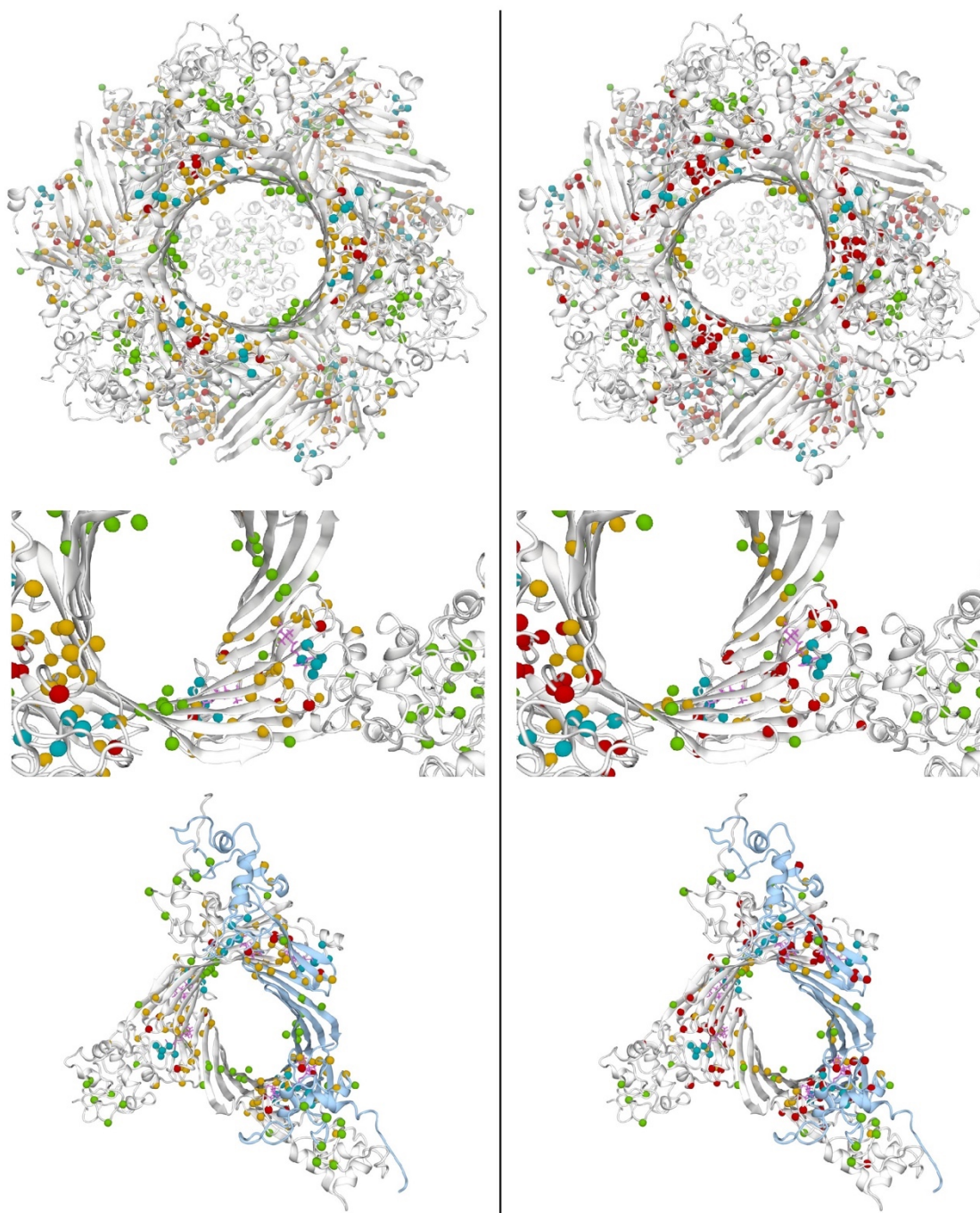
Figure 23: Signal bleaching effect of MTSL (left) and TOTAPOL-derivative (right) mapped on the <u>Weinkauf</u>**-model**[77] **(PDB-entry 2YGD). The cyan-colored balls mark the stretch 158 – TICIT – 162. Residue P160 was mutated in PyMol to cysteine and the label (MTSL, pink) manually added in Autodesk Maya. For convenience, MTSL is shown in both panels to the left and to the right, although the PRE-effect of the TOTAPOL-derivative is visualized on the right. Red residues experience a strong effect, orange residues an intermediate effect and green residues no PRE-effect. For blue or whitish areas quantification of the PRE-effect through the paramagnetic label was not possible due to signal-overlay in the spectra or missing assignment. Pictures by Barth-Jan van Rossum.**
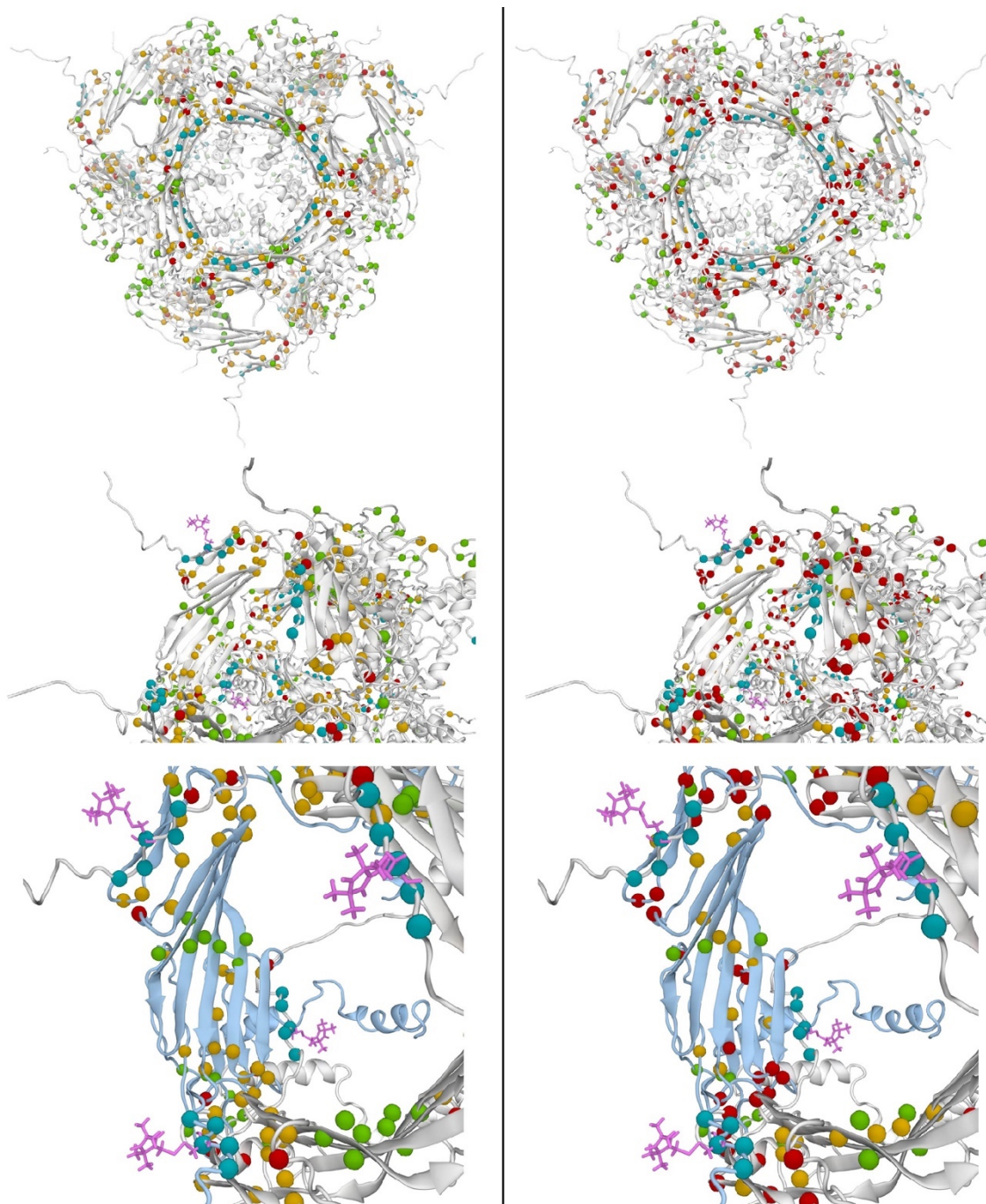
Figure 24: Signal bleaching effect of MTSL (left) and TOTAPOL-derivative (right) mapped on the <u>Jehle</u>-model[68] (PDB-entry 3J07). The cyan-colored balls mark the stretch 158 – TICIT – 162. Residue P160 was mutated in PyMol to cysteine and the label (MTSL, pink) manually added in Autodesk Maya. For convenience, MTSL is shown in both panels to the left and to the right, although the PRE-effect of the TOTAPOL-derivative is visualized on the right. Red residues experience a strong effect, orange residues an intermediate effect and green residues no PRE-effect. For blue or whitish areas quantification of the PRE-effect through the paramagnetic label was not possible due to signal-overlay in the spectra or missing assignment. Pictures by Barth-Jan van Rossum.

3.4 Docking-model of the IXI-motifs

As was shown by X-ray crystallography[70] and solid-state NMR-spectroscopy[68], the C-terminal IXI-motif can bind to the hydrophobic groove formed by β4 and β8. The NMR-restraints revealed contacts for amides of S135 (β8) with Cγ2 of I159, K90 (β4) with Cδ of I161, L137 (β8) with Cγ of T162 and I161 with Cγ2 of V91 (β4).

The sequence stretch 156 – ERTIPITRE – 164 around the C-terminal IXI-motif is palindromic and found to bind to the hydrophobic groove in both orientations in binding studies using a short construct αB-crystallin 68-162[70]. However, due to the relatively low resolution of the solid-state NMR-data of the full construct, it was not possible determine a second set of restraint that indicated a possible binding in both directions.

From the current study we now know that the signals of I3 and I5 overlap with the ones of I159 and I161 in the solid-state NMR data; hence, the contacts detected by Jehle could equally well be assigned to the interaction of the N-terminus with the groove.

Even though the neighboring residues around the two IXI-motifs are different, with 1 – MDIAIHH – 7 for the N- and 155 – PERTIPITR – 163 for the stretch around the C-terminus, both sequence stretches can be docked into the hydrophobic binding groove (**Figure 25)**. A similar binding of the N-terminal motif to the hydrophobic groove was found in case of the small heat shock protein Tsp36 from *Taenia saginata,* a flatworm species[151]. Tsp36 is lacking the particular motif in the C-terminus but possesses two α-core domains. The crystal structure of Tsp36 (PDB-entry 2BOL) revealed that the N-terminal IXI-motif binds to the hydrophobic groove formed by β-strands 4 and 8 of the ACD1 of a neighboring monomer.
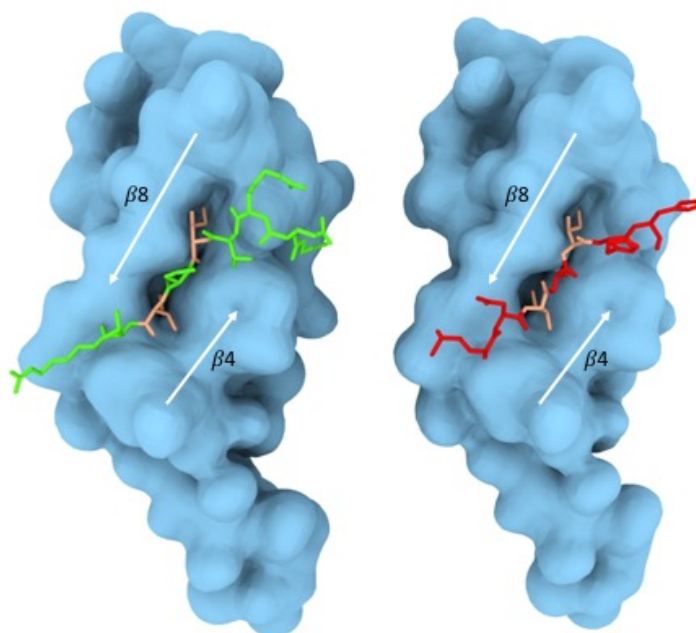
Figure 25: Docking model of N- and C-terminal IXI-motifs to the hydrophobic groove formed by β4 and β8, indicated with arrows. N-terminal sequence M1-D2-I3-A4-I5-H6-H7 (red, right) and C-terminal sequence P155-E156-R157-T158-I159-P160-I161-T162-R163 (green, left) displayed as stick models on surface-model of the α-crystallin domain (light blue). The isoleucines of the IXI motives are highlighted in orange Picture made by Barth-Jan van Rossum.

3.5 The results in context to other published data on αB-crystallin

Poulain *et al.* investigated the structural and sequential preservation among known α-crystallin domain structures and found that even though the sequence similarity is rather low, seven known structures for ACDs (αB-crystallin, HSP16.5, HSP A, Tsp36.1, HSP16.9, HSP20, Tsp36.2) overlay almost completely[81]. This finding shows that the core domain is structurally (but not sequentially) highly conserved and emphasizes a critical structure-functional role of the core-domain in αB-crystallin, which is related to the chaperoning capabilities towards certain targets[74]. Similarly, subsequences of high sequential conservation, such as the IXI-motif can be assumed to be functionally relevant units. Given its high degree of conservation in the CTDs of small heat shock proteins in general (96% as L/V/I – X – I/V/L) and in the NTDs among human homologues, it is tempting to assume a similar function for both motifs. The interaction of this motif with the hydrophobic groove and its role in auto-inhibition by covering it can provide an essential mechanism for regulation of chaperone activity. Such mechanism was demonstrated by Pasta *et al.* who showed

the structural and functional importance of the motif in temperature-dependent chaperoning[46]. With two IXI motifs, αA-crystallin, αB-crystallin and HSPB9 are theoretically equipped with two mechanistically equivalent features for self-control of chaperoning via binding of the IXI-motifs to the hydrophobic groove. However, it remains unclear if the function of the IXI motifs is the same in case of αA and HSPB9. In the case of αA, the oligomerization behavior is found to depend on the arginine after the C-terminal IXI-motif (156 – ERAIPVS<u>R</u>E – 164), as the particle size was altered significantly in a short construct that was cut at the serine preceding the arginine[161]. A similar dependency, however, is not known for αB-crystallin, as, to my knowledge, short constructs were always truncated at both termini. On the other hand, investigations on spin-labelled αB-crystallin at position 135 and 146 revealed that the interdimer contacts derive from residues upstream the IXI-motif[162].

The results from this work shed new light on the finding about the proteins behavior upon pH-drop published by Jehle *et al.*[68]. SAXS-measurements of samples at pH 7.5 and pH 6.5 revealed an increase of gyromagnetic ratio R_G and particle diameter D_{max}. Together with the observation of chemical shift changes in the isoleucine region in PDSD spectra (**Figure 26**), this was interpreted as activation of the oligomer by unbinding the C-termini. The presence of additional free C-terminus was used to explain the increased particle size and assumed to result in activation of the oligomer under pH-induced cellular stress conditions.

Early results on αB-crystallin revealed a complementary size-behavior under salt-stress or conditions of elevated temperatures[54] and it was argued that smaller particles lead to a higher surface/volume ratio and therefore to more exposed binding site. This is in direct contradiction to the pH-drop results. With the current knowledge that the N-terminus can as well bind to the hydrophobic groove and taking into consideration that the IXI-motifs of both the N- and the C-termini have very similar chemical shifts for the isoleucines, one could now reconsider the previously reported shift of I159 upon pH-drop. In retrospect, the change of the Cγ1δ cross peak of I159 could as well be attributed to an N-terminal isoleucine. The explanation for the pH-dependent increase of particle size could lie in the hydrophilic properties of the CTD and its contribution to the solubility of the protein.
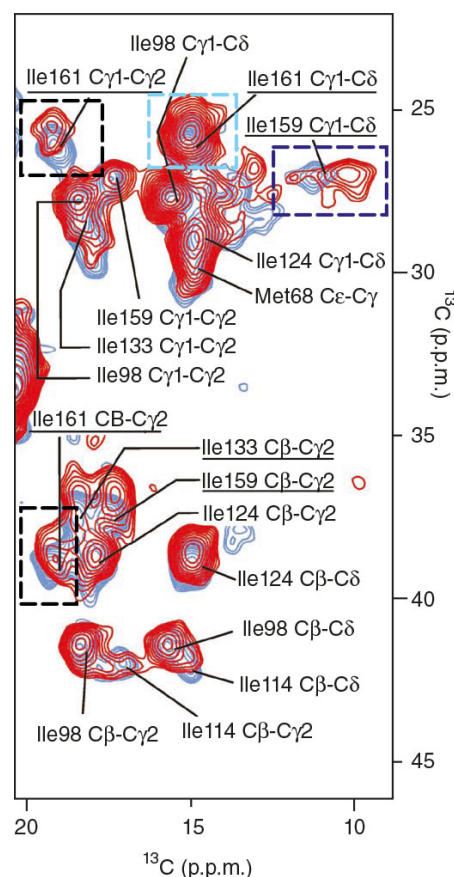
Figure 26: Plot of αB-crystallin WT at pH 7.5 (blue) and pH 6.5 (red) from Jehle *et al.*[68]. Dashed boxes indicate the shifting isoleucine signals, which was interpreted as detachment of the C-terminal IXI-motif from the hydrophobic groove.

The source for the second set of signals in the HSQC-spectra of αB-crystallin and mutants remains unknown. As Mainz pointed out, the major and minor species experience exchange on a minute time-scale, which would correspond to either subunit exchange or proline cis-trans-isomery. In the ^{13}C-direct excitation data this second set of signals was not observed, which can be due to a lower resolution in the 1D data. However, if the minor set in the HSQC is due to cis-trans-isomery, it could well be that its effect does not affect long sidechains and therefore could escape detection in the 1D data. Furthermore, the question would remain which prolines exhibit the isomeric behavior, if it is only one or more of the C-terminal prolines and how far the effect on the chemical shifts of neighboring residues reaches. In case of the αB P160C mutant, we observe a second set for I159 but not for I161. In contrast, for the wild-type, a minor species is found for both motif-related residues. The effect at position 159 could also derive from isomerization of the preceding P155.

Investigations on Hsp27 revealed cis-trans-isomery on P194[163], so in case of αB-crystallin, cis-trans isomery on one or more proline residues of the CTD cannot be ruled out so far.

A further matter of controversy is raised by our results on the spin-label loss. Investigations of αB I5C and αB T162C under oxidative conditions by Murugesan *et al.* revealed that covalent dimer formation was three times higher for the C-terminal cysteine-mutant[164]. Our mutants carry the cysteines at very similar positions but the results from the EPR-measurements show a contrary behavior as the A4C-variant loses the spin-label very fast.

4 Conclusion and perspectives

The degree of structural and sequential preservation in the small heat shock protein-family underline the importance of certain features. This includes the structural preservation of the α-crystallin domain[81], and the sequential conservation of certain residues or motifs in all three domains of the protein[13], among others the terminal IXI-motifs. This work demonstrated that among the human homologues not only the C-terminal IXI-motif belongs to the highly conserved sequences in the sHSP-family, but that the N-terminal IXI-motif is preserved as well.

When Augusteyn suggested the micelle-model, he criticized all other published models because of the failure to explain the distribution of oligomers. He argued that the models exhibiting tetrahedral symmetry and structure with different layers should be reflected in the oligomeric distribution by a higher occurrence of particles where the subunit number can be divided by 12[55]. The same problem of real size-distributions not reflecting the expected distribution for the proposed models arises if we emphasize the hierarchy of interactions for oligomer built-up too strongly instead of regarding the interactions as mechanisms that can take place independently from each other. Hence, the argumentation from Augusteyn still holds true for the currently much better understood interactions leading to oligomers. For instance, if dimer formation is considered as the first order interaction, then even-numbered oligomers should be more emphasized in the distribution of sizes, which is only partly true according to results from MS[48]. It seems appropriate to take a step away from the hierarchical picture of oligomer formation in a set of distinct stages, but rather switch in the prospective description of the protein to interactions that take place simultaneously rather that sequentially, without recurring to attribute them as "levels of interaction".

In the history of proposed models for α-crystallin, and later on for αB-crystallin, the endeavor to come up with a symmetrical structure was always strong. If the function of the protein is to be a chaperone upon request that readily exposes its binding sites on the surface[62,74], natures' strive for symmetry might not be essential for the

73

function, but rather a feature which comes along to a certain extend. It is more likely that the statistic equidistant distribution of binding sites along the surface of the spherical particle is essential for the best possible function as a chaperone.

In this work we were interested to provide a better understanding of the function of the protein-termini in the oligomeric ensemble with particular interest in the role of the IPI-motif of the C- and IAI-motif in the N-terminus. We showed that the N-terminal IAI-motif is competing with the C-terminus for the binding to the hydrophobic groove.

By mutation of the C-terminal isoleucines into valines, the N-terminal motif-related isoleucines were made visible in the 2D solid-state NMR-data. This allowed us to assign both the N- and C-terminal isoleucine motifs for the first time. Interestingly, it turned out that the chemical shifts of the isoleucines of both motifs overlap to a high extend.

The ^{13}C-direct excitation spectra made the rigid conformations and flexible parts of the protein visible in the same spectra for the first time. By subtracting the spectra of different valine-mutants, we could determine the portions of bound and unbound IXI-motifs and show the temperature-dependency of the occupation of the hydrophobic groove in relation to the free motif.

The PRE-effects in the spin-labelled samples and the results from the EPR and DEER-measurements underline the pronounced inherent heterogeneity of αB-crystallin. DEER-measurements illustrated the sheer endless possible inter-IPI-motif distances within the oligomers. In case of the biradical, the PRE-effect is extended from the loops connecting the β-strands of the core to the barrel-like structure of the α-crystallin domain. The N-terminal residues which are not affected by the PRE-effect of both radicals are surface exposed in the Jehle-model and embedded in the center of the N-terminal bridge-heads in the Weinkauf-model. Given the overall decrease of signal intensity, embedment could be the more likely state for residues experiencing no recognizable PRE-effect. Both models fail to explain the partially strong PRE-effect on A57 and the lacking bleaching of the cross peak for S136.

The fact that the spin-label cleaves off very quickly in the αB A4C variant can be interpreted as the cysteines being located in close proximity to each other in the

oligomer, a situation that would be in line with both the Jehle- and the Weinkauf-model.

We conclude that both termini have multiple roles regarding structure and function of the sHSP. The C-terminus contributes to the oligomerization-process in its bound state and provides with its high polarity to the excellent solubility of the protein in its flexible state. The N-terminus provides to the formation of higher order oligomers as determined through reverse conclusion of the oligomerization behavior of short constructs[70]. This happens through hydrophobic interactions with other NTDs and through binding of the IAI-motif if this is a heteromeric interaction. In addition, it is essential for the chaperone function towards certain targets[19,74,95,101,165].

The ability of the IPI-motif to cover the hydrophobic groove is believed to bear a regulatory function for chaperoning[68]. In this work we found that the N-terminal IAI-motif provides a similar functionality. This increases the possibilities not only for regulation but also for structural plasticity. It is very important to realize that both 24mer models only provide a single, idealized arrangement for a particle comprising exactly 24 subunits, but do not provide a general solution for building oligomers from an arbitrary number of subunits. Having two IXI-motifs that are able to cover the hydrophobic groove increases the possibilities to construct oligomers from a number of subunits different than 24, that still incorporate protective mechanisms for self-inhibition to prevent the protein of overreaction under normal cellular conditions and furthermore increase the heterogeneity.

Not every structural detail in oligomer-formation is known or fully understood. Yet the results from this work provide an important new aspect concerning the role of the N-terminus of αB-crystallin. Its contribution to the inherent heterogeneity of the protein is in line with the diversity of interactions on the intra- and intermolecular level which results in a complexity and polydispersity which is hard to investigate in every detail but leads in the consequence to a perfect adaption of the protein to the circumstances and its tasks and allows a fast chaperone-reaction in cases of cellular stress.

Taking into account the number of known different possible combinations by which the αB-crystallin monomers can interact with each other, the inherent heterogeneity and polydispersity of the protein comes clear

The work and results in this Thesis provide a basis that might motivate or stimulate a series of future experiments. First of all, to circumvent the problem of label-loss like it happened with the radical-label at the cysteine in the αB A4C, , protein-observed fluorine NMR[166] could serve as alternative. Fluorine does not require isotope labelling as it is 100% natural abundant as ^{19}F and a very NMR-active element. Since αB-crystallin contains only two tryptophanes, at positions 10 and 60, NMR-spectra of αB labelled with fluorinated tryptophane, should be sparsely populated and reveal very specific knowledge about the W10. Label-loss would not be an issue as the fluorine-carbon bond is very stable. In combination with a possible point mutation of residue W60 the amount of different states for the very beginning of the N-terminal domain would be assessable.

The activation of the protein upon pH-drop as published by Jehle *et al.*[68] could be investigated under the new light of the IAI-motif signal pattern in the solid-state NMR data in this work. The pH-drop performed on samples for ^{13}C-direct excitation could possibly reveal interesting details on the ratio of the binding of the two IXI-motifs in a pH-dependent manner.

To further probe the effect of the N-terminal IAI-motif on the binding to the groove, investigations on motif-related mutated versions of the protein or short constructs lacking the first five or more residues, could be useful. This could also provide additional information about the bridge-head-forming features and possibly reveal if the N-terminal motif-binding is a homomeric or heteromeric event. Mutations could be considered that sterically inhibit the binding to the hydrophobic groove, or a short construct could be created that misses the lowest possible number of residues in order to maintain other functional properties provided by the N-terminus.

In all experiments on αB-crystallin, being part of this thesis or future work, the inherent heterogeneity is and will be a challenge. Nevertheless, the importance of the protein for the homeostasis in cells, underlines the need for finding answers to questions about the structure and activation mechanisms of the sHSP.

Appendix
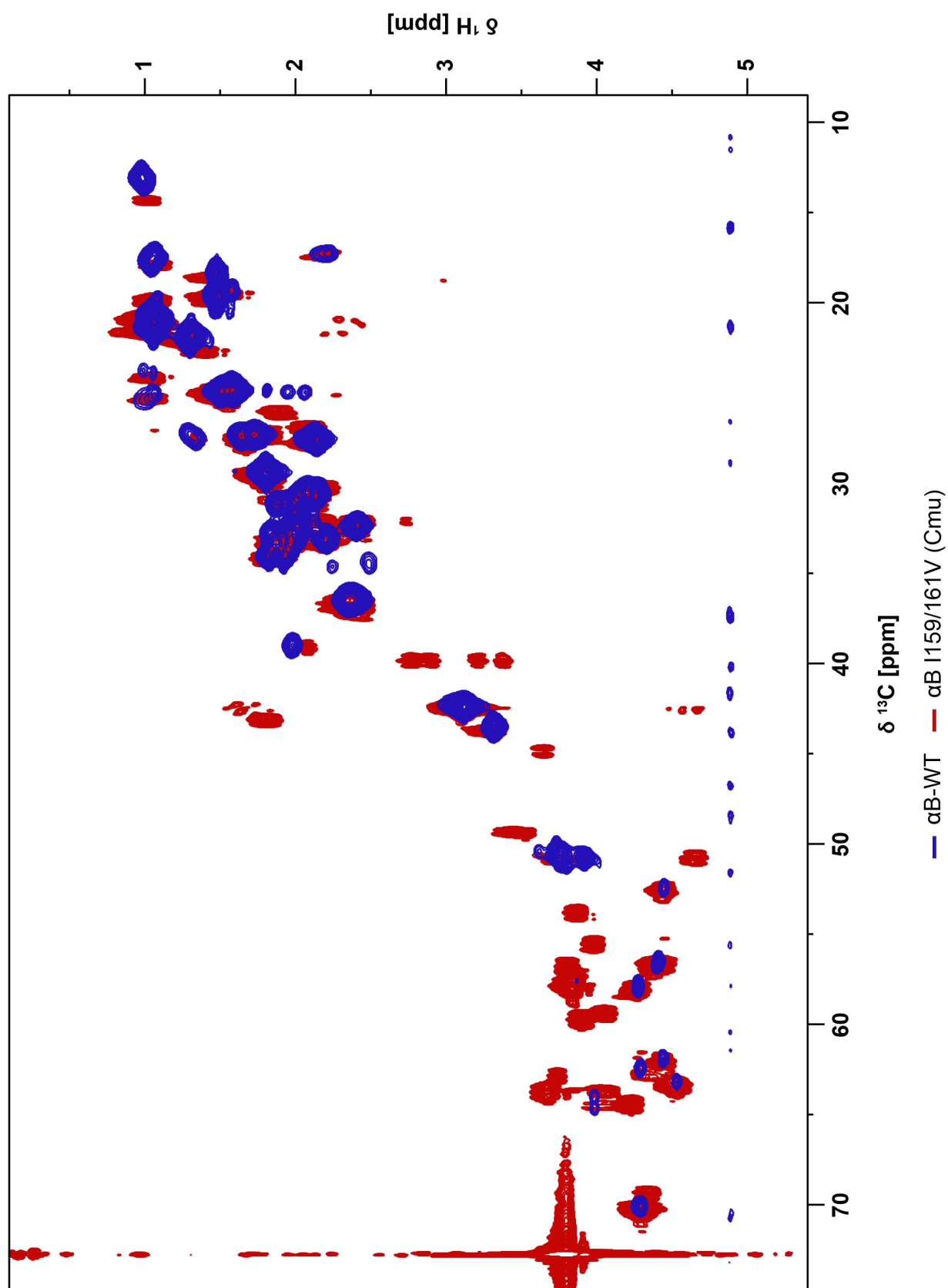
Figure 27: Overlay of solution NMR CH-correlation of αB-WT (dark blue) on MAS solid-state NMR HC-INEPT of αB I159/161V (red).
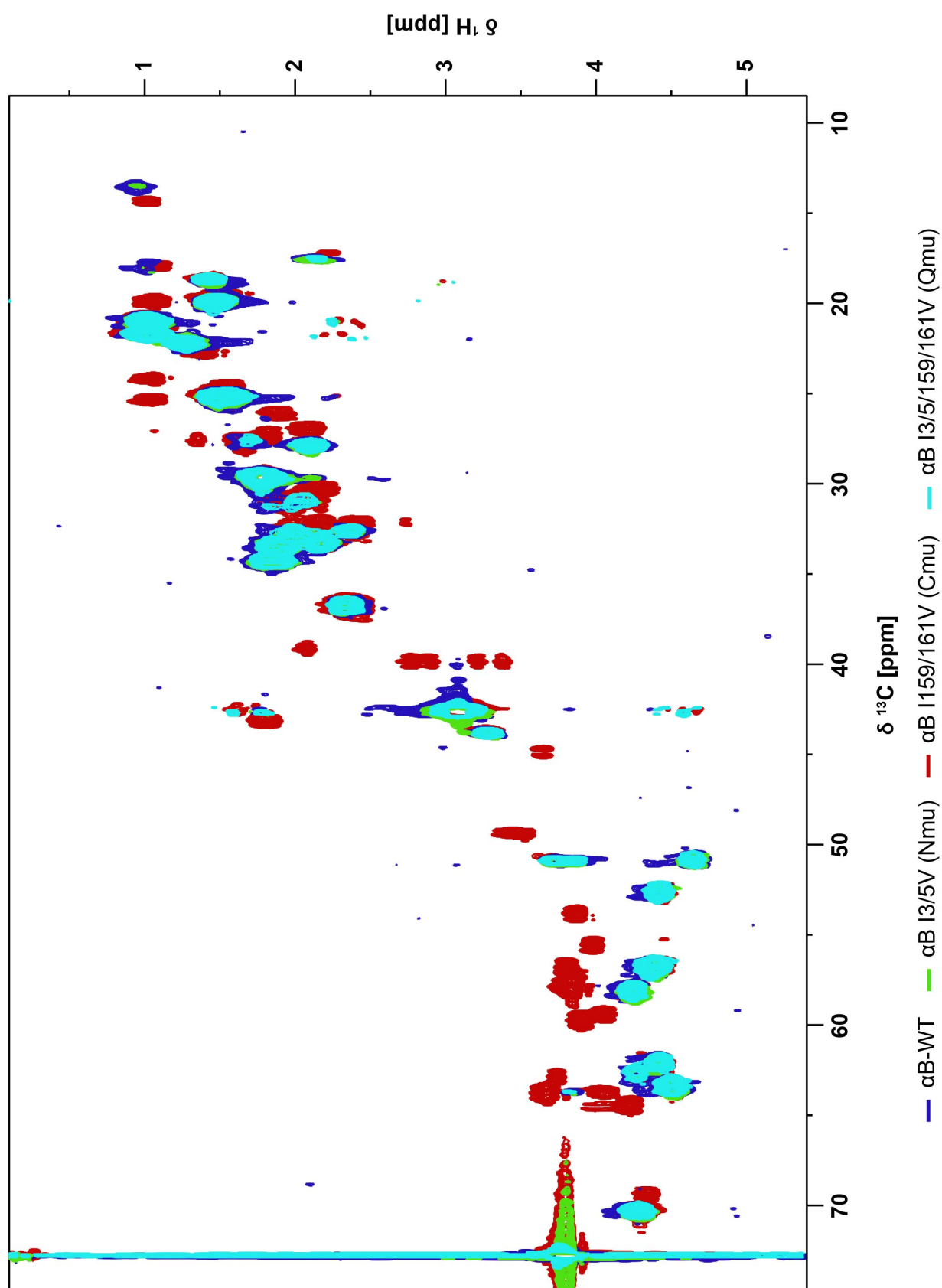
Figure 28: MAS solid-state NMR HC-INEPT spectra of αB-WT (dark blue) and the valine mutants αB I3/5/159/161V (light blue), αB I3/5V (green) and αB I159/161V (red) recorded from PEG-precipitates.
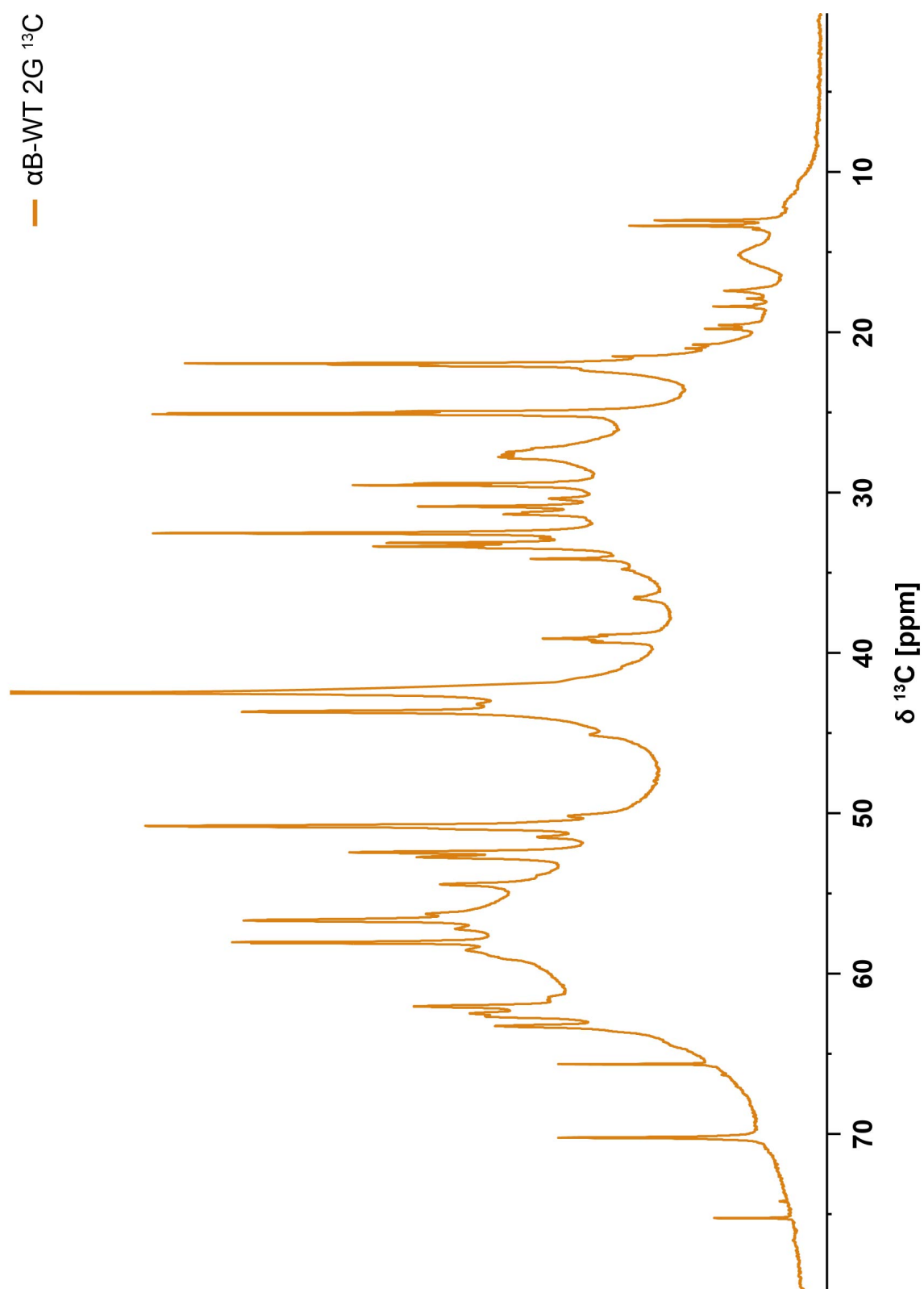
Figure 29: ^{13}C-direct excitation spectrum of 2G-^{13}C labelled αB-crystallin recorded at 290 K. The delay-time before every scan was set to 30 s in order to assure complete relaxation between experiments.
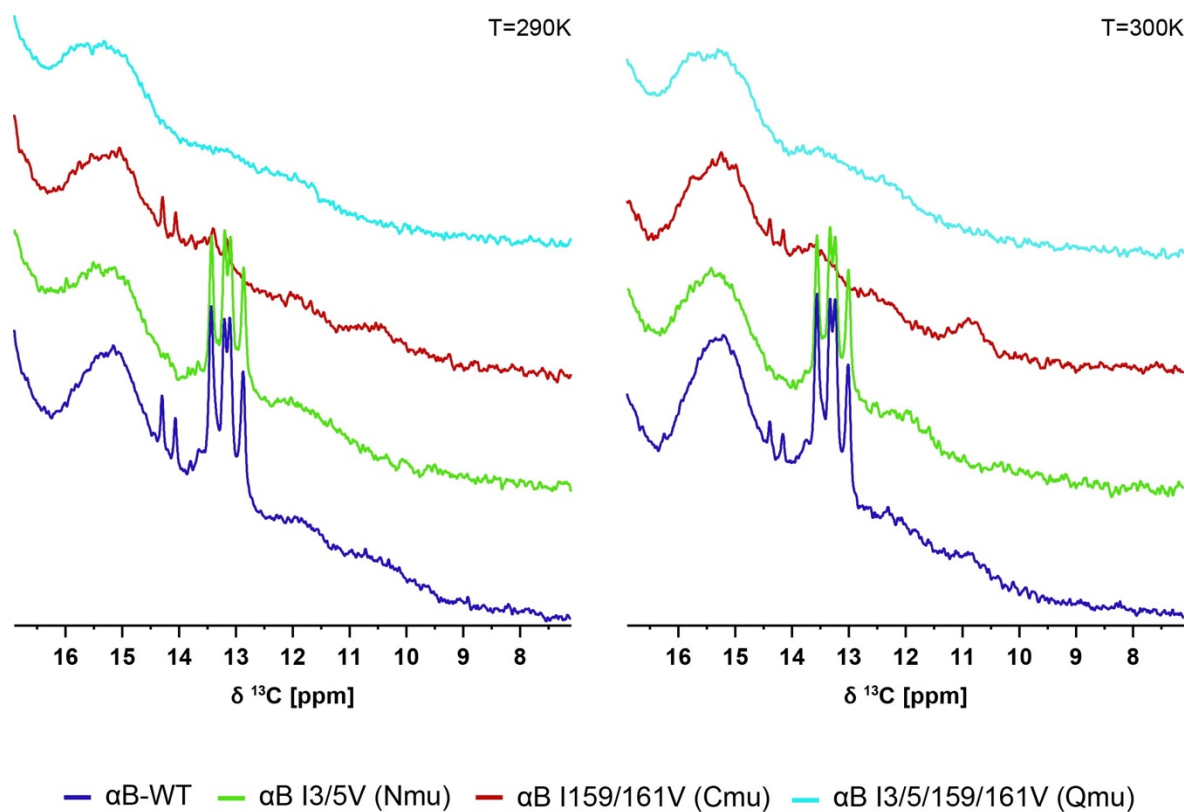
Figure 30: Upfield aliphatic range of ^{13}**C-direct excitation spectra of** ^{13}**C-labelled αB-WT and valine mutants, recorded at 290 K (left) and 300 K (right). Color coding for the various mutants is indicated in the figure. The delay-time before every scan was set to 30 s in order to assure complete relaxation between experiments.**
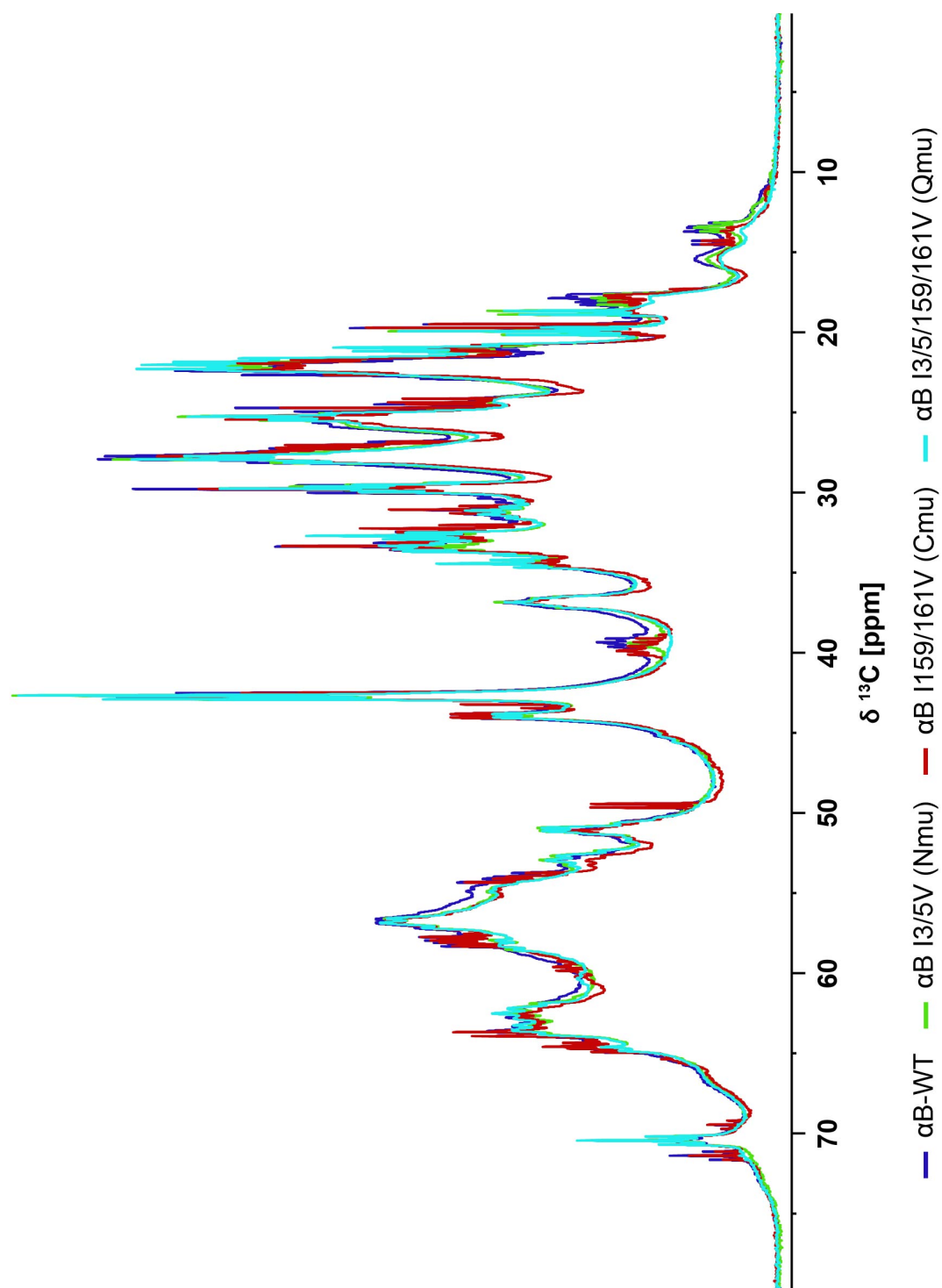
Figure 31: ^{13}C-direct excitation spectra of ^{13}C-labelled αB-WT and valine mutants, recorded at 310 K. Color coding for the various mutants is indicated in the figure. The delay-time before every scan was set to 30 s in order to allow for sufficient relaxation between the experiments.
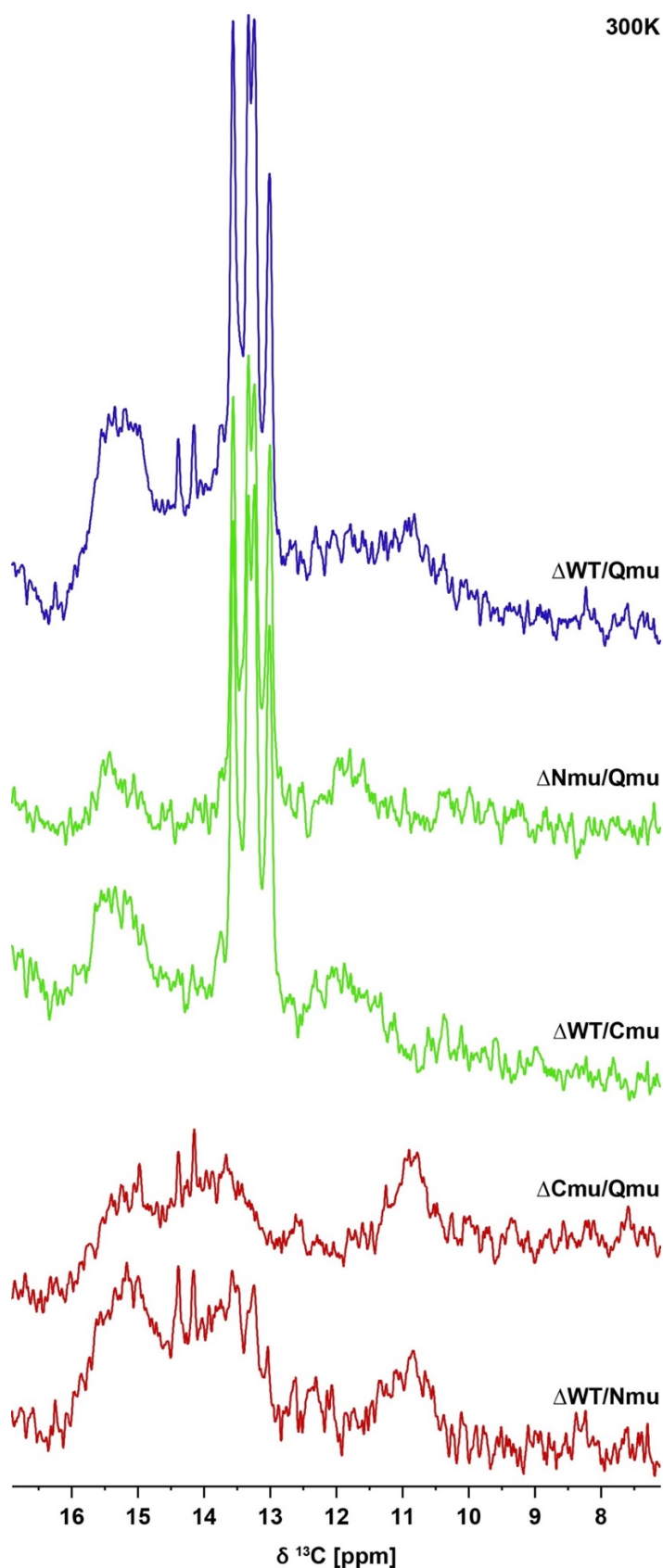
Figure 32: Difference-spectra obtained by subtraction of respective ^{13}C direct-excitation spectra recorded of the various mutants at 300 K. Color coding refers to the remaining motif-related positive isoleucine-signals resulting in dark blue for ΔWT/Qmu as all four motif-related isoleucines give positive signals, light green for positive C-terminal isoleucines, red for positive N-terminal residues.
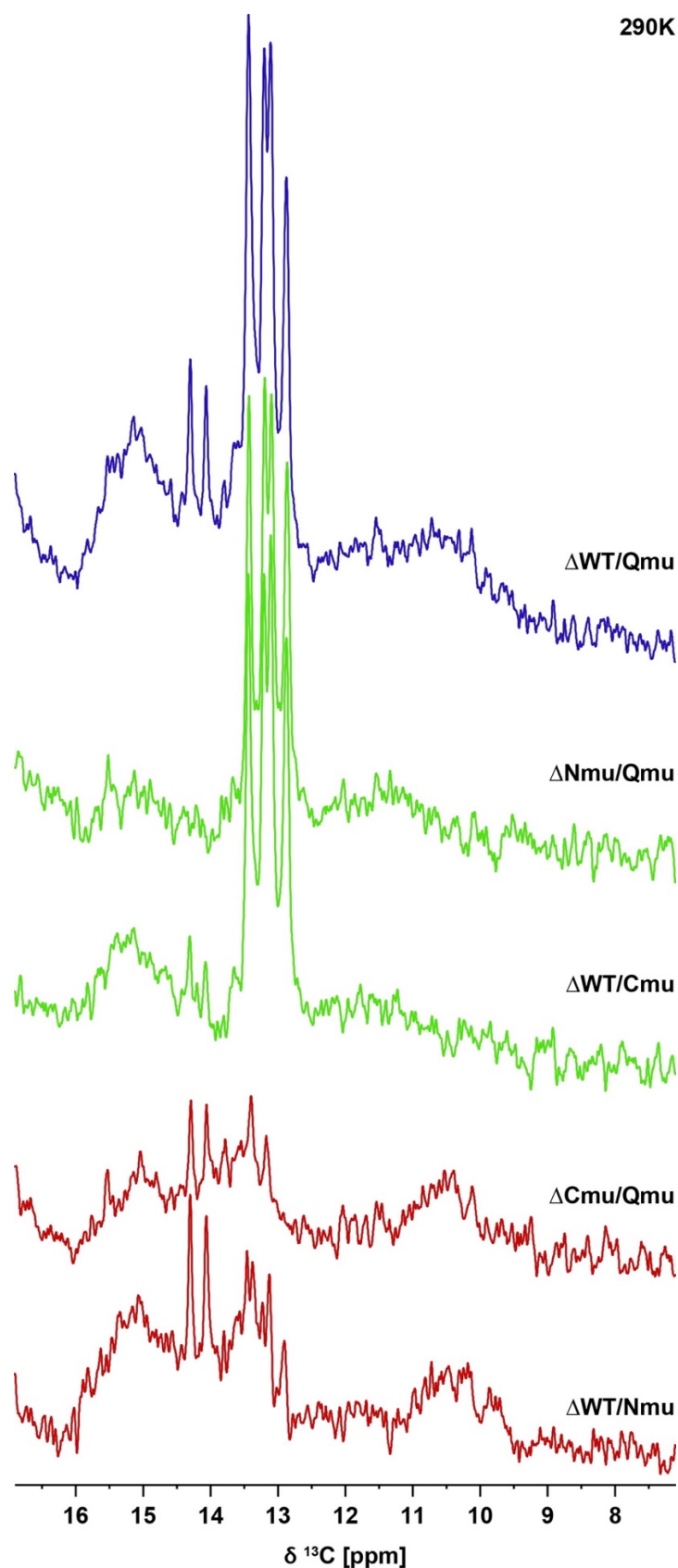
Figure 33: Difference-spectra obtained by subtraction of respective ^{13}C direct-excitation spectra recorded of the various mutants at 290 K. Color coding refers to the remaining motif-related positive isoleucine-signals resulting in dark blue for ΔWT/Qmu as all four motif-related isoleucines give positive signals, light green for positive C-terminal isoleucines, red for positive N-terminal residues.

Table 9: Integration areas in ppm for difference-spectra (Figure 19, Figure 32 and Figure 33). Obtained values are listed in Table 8.

substracted spectra	T[K]	unbound Alanine (A4)	unbound Isoleucines (I3/5/I159/161)	bound C-terminal Isoleucine (I159)	bound N-terminal Isoleucine (I3 or I5)
ΔWT/Nmu		14,40-13,96	13,85-12,85	-	11,08-9,53
ΔWT/Cmu	290	-	13,54-12,75	12,02-10,62	-
ΔWT/Qmu		14,41-13,95	13,92-12,73	12,02-9,53	
ΔWT/Nmu		14,52-14,04	14,03-12,96	-	11,53-10,31
ΔWT/Cmu	300	-	13,69-12,90	12,52-10,94	-
ΔWT/Qmu		14,54-14,03	14,03-12,88	12,52-10,31	
ΔWT/Nmu		14,60-14,17	14,17-13,03	-	11,97-10,69
ΔWT/Cmu	310	-	13,86-13,02	12,60-11,33	-
ΔWT/Qmu		14,60-14,13	14,13-12,99	12,60-10,69	
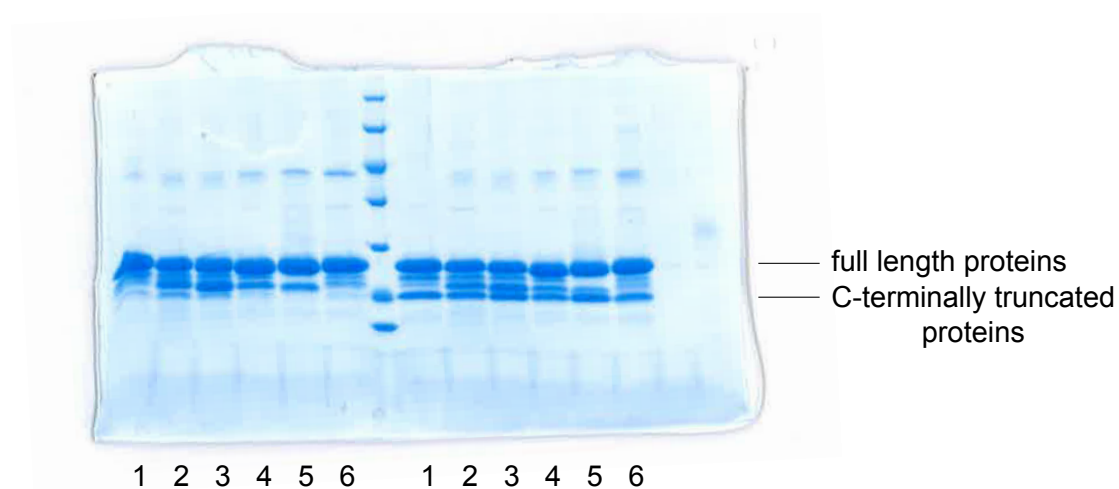
Figure 34: Proteolytic cleavage of the C-termini in αB-crystallin and mutants with enterokinase as shown by Mainz *et al.*[74]. Reaction time was 40 h at 4 °C. Bands on the left from protein-solutions before the proteolysis, bands to the right after cleavage of C-terminal residues 158 to 175, reference in the middle. Bands 1 to 6 from proteins as follows: 1) αB WT, 2) αB I3/5V, 3) αB I159/161V, 4) αB I3/5/159/161V, 5) αB P160C and 6) αB A4C.

A) DEER-traces with fitting and curves for distances for A4C-MTSL
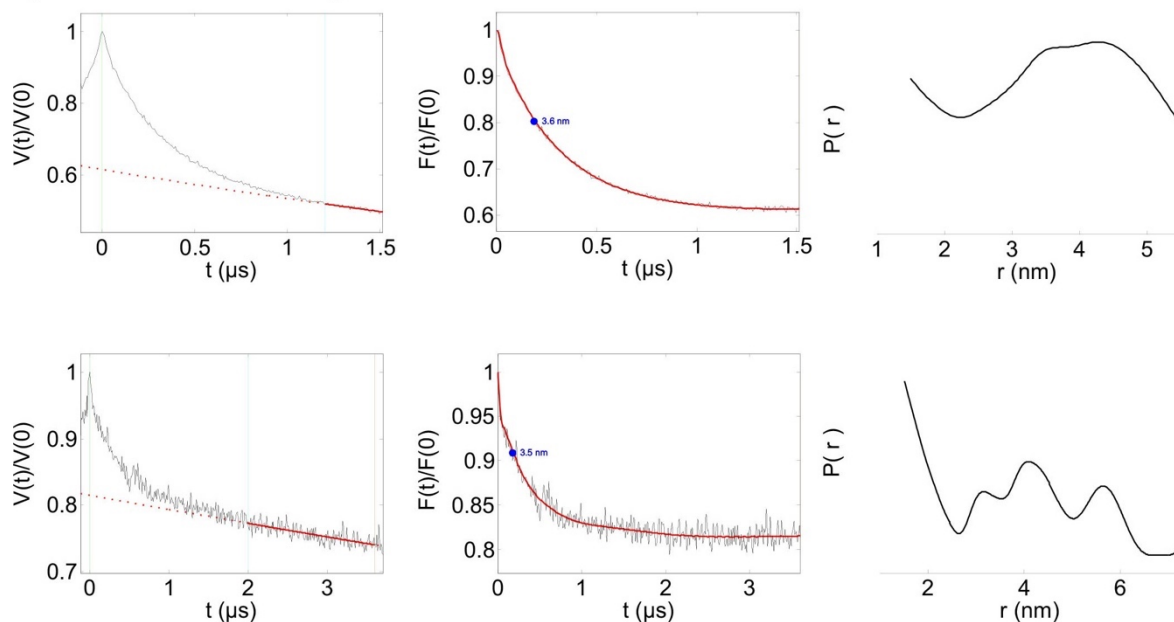
B) DEER-traces with fitting and curves for distances for P160C-MTSL
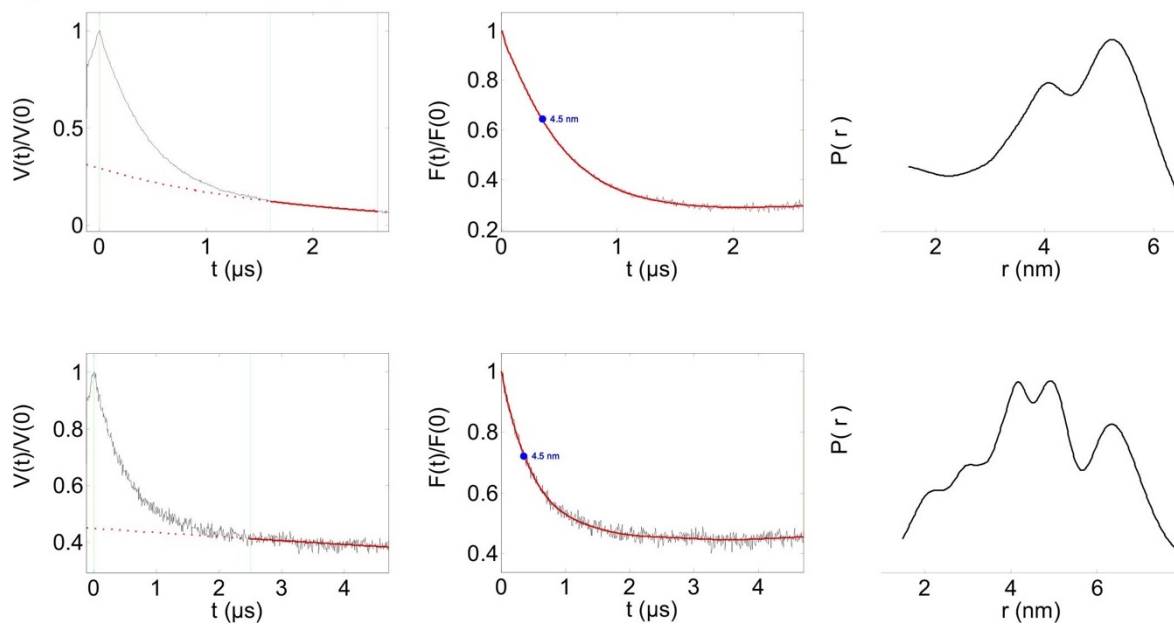
Figure 35: DEER-traces and label-distance curves for MTSL spin-labelled A4C and P160C for two different acquisition times for each sample.
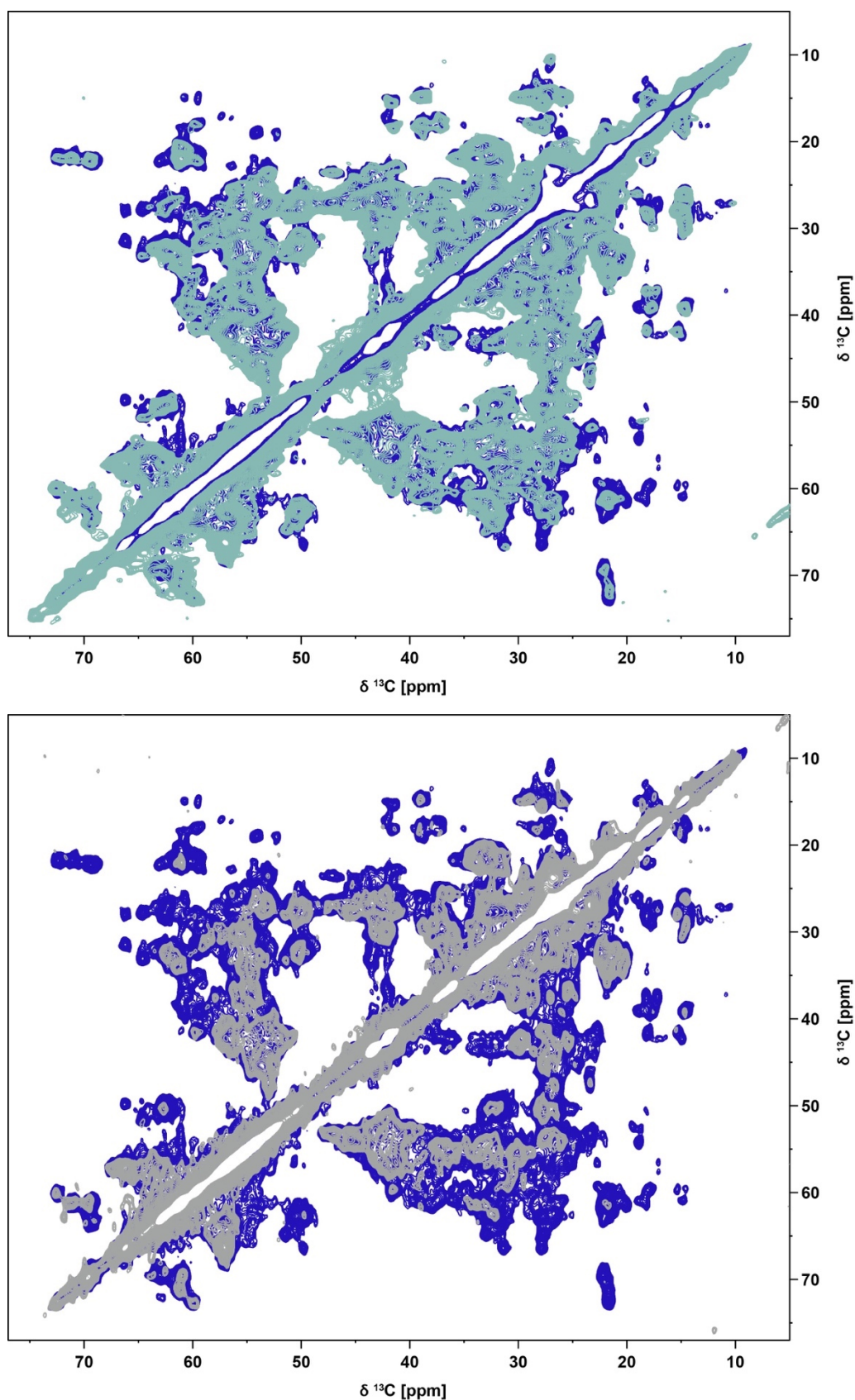
Figure 36: Sidechain area of ^{13}C-^{13}C PDSD-spectra of spin-labelled αB P160C-MTSL (turquoise, top) and αB P160C-TOTAPOL (light grey, bottom) plotted on top of αB-WT spectrum (dark blue). All spectra were recorded with similar parameters, for details see Materials & Methods section chapter 2.3.2 Table 5.

Table 10: Summary of categorization of spins for experienced PRE-effect in αB-P160C labelled with either MTSL or TOTAPOL-derivative SB-II-72. Categorization carried out according to strength of PRE-effect in 2D PDSD spectra on precipitated protein-samples compared to the wild-type spectrum. Red for strong effect, orange for affected but not highly affected spins, green for not affected ones, blue for not assigned or not resolved spins, the underlying assignment is taken from Jehle[67] (*BMRB* Entry 16391).

MTSL-label	*red*	86, 126, 133, 159, 161
	orange	57, 64, 77, 79, 89, 95, 97, 98, 124, 129, 130, 140, 141, 147, 162
	green	40, 42, 45, 68, 75, 102-104, 117
	blue	1-39, 41, 43, 44, 46-56, 58-63, 65-67, 69-74, 76, 78, 80-85, 87-94, 96, 99-101, 105-116, 118-123, 125, 127, 128, 131, 132, 134-139, 142-146, 148-158, 160, 163, 164, 165-175
TOTAPOL-derivative-label	*red*	57, 77, 86, 95, 126, 129, 130, 133, 140, 141, 147, 159, 161, 162
	orange	64, 79, 98, 102, 117, 124, 136
	green	42, 45, 68, 75, 103, 104
	blue	1-41, 43, 44, 46-56, 58-63, 65-67, 69-74, 76, 78, 80-85, 87-94, 96, 97, 99-101, 105-116, 118-123, 125, 127, 128, 131, 132, 134, 135, 137-139, 142-146, 148-158, 160, 163-175

Table 11: Chemical shift list for unambiguously assigned resonances for αB I159/161V in solution NMR-spectra. For type of spectra recorded for assignments, see Table 4. The assignment was accomplished in *Sparky*.

Residue	H_N	N_H	Cα	Cβ	others
S153	-	-	58.50	64.46	-
P155	-	-	63.33	32.30	-
E156	8.42	120.81	56.98	30.08	-
R157	8.14	121.89	56.23	31.05	Cγ: 27.22, Cε: 43.52
T158	-	-	61.98	70.28	-
V169	8.01	124.01	60.01	32.95	-
P160	-	-	63.10	32.20	-
V161	8.09	120.60	62.47	33.19	-
R163	-	-	56.26	31.16	Cγ: 27.46, Cε: 43.71
E164	8.28	122.51	56.52	30.66	-
E165	8.30	123.20	56.45	30.54	-
K166	8.23	123.82	54.27	32.59	-
K166b	7.97	121.96	-	-	-
P167	-	-	63.11	32.48	-
A168	8.25	124.63	52.57	19.41	-
V169	8.01	119.50	62.45	33.18	Cγ1: 21.49, Cγ2: 20.82
V169b	8.14	120.29	62.65	-	-
A171	8.10	126.88	52.27	19.53	-
A172	8.08	125.01	50.63	18.14	-
A172b	7.86	123.07	-	-	-
P173	-	-	63.12	32.48	Cγ: 27.67, Cδ: 50.78
K174	8.21	122.24	56.51	33.07	Cε: 42.41
K175	7.79	128.06	57.88	33.87	Cγ: 24.85, Cδ: 29.48, Cε: 42.43
K175b	7.83	128.41	-	-	-

Table 12: Chemical shift list for unambiguously assigned resonances for αB-WT in solution NMR-spectra. For type of spectra recorded for assignments, see Table 4. The assignment was accomplished in *Sparky*.

Residue	H_N	N_H	C	Cα	Cβ	others
S153	-	-	175.44	58.49	64.41	-
G154	-	-	172.97	44.87	-	-
E156	8.44	120.76	-	-	-	-
R157	8.14	121.91	177.19	56.18	31.09	Cγ: 27.20
I159	-	-	175.42	58.80	38.98	Cγ1: 27.32, Cγ2: 17.42, Cδ: 13.00
I161	-	-	177.55	61.61	39.20	Cγ1: 27.64, Cγ2: 17.84, Cδ: 13.34
R163	-	-	176.82	56.18	31.25	Cγ: 27.36
E165	8.30	123.12	-	-	-	-
K166	8.23	123.74	175.25	54.36	32.84	Cγ: 24.86, Cδ: 29.45, Cε: 42.38
K166b	7.97	121.96	-	-	-	-
A168	8.25	124.64	178.94	52.60	19.38	-
V169	7.99	119.42	177.39	62.55	33.19	Cγ1: 21.47, Cγ2: 20.75
A171	-	-	177.83	52.27	19.69	-
A172	8.09	125.02	176.49	50.67	18.36	-
P173	-	-	177.80	63.14	32.41	Cγ: 27.68, Cδ: 50.80
K174	8.21	122.25	176.77	56.64	33.23	Cγ: 24.84, Cδ: 29.36, Cε: 42.44
K175	7.79	128.03	182.24	57.93	34.09	Cγ: 24.94, Cδ: 29.45, Cε: 42.45

Table 13: Chemical shift list for unambiguously assigned resonances for αB P160C in solution NMR-spectra. For type of spectra recorded for assignments, see Table 4. The assignment was accomplished in *Sparky*.

Residue	H_N	N_H	C	Cα	Cβ
P155	-	-	-	63.54	31.69
E156	8.44	120.73	177.60	57.08	29.76
R157	8.14	121.81	-	56.12	30.83
T158	8.12	118.90	175.37	62.21	69.73
T158b	-	-	-	62.36	69.73
I159	8.03	123.43	-	61.25	38.52
I159b	8.09	123.49	-	61.23	38.50
C160	8.23	119.67	175.38	58.55	27.73
I161	8.15	124.32	176.68	61.42	38.48
T162	7.63	123.42	-	63.37	70.87
R163	-	-	176.87	56.32	30.56
E164	8.27	122.50	177.17	56.49	30.04
E165	8.30	123.19	176.67	56.51	30.22
E165b	8.43	123.42	-	56.44	30.12
K166	8.24	123.81	-	54.16	32.21
K166b	7.98	122.02	-	53.78	33.99
P167	-	-	177.67	63.13	31.80
P167b	-	-	177.59	62.99	31.94
A168	8.26	124.60	178.96	52.50	18.87
A168b	8.37	124.72	178.97	52.70	18.91
V169	8.00	119.51	-	62.41	32.48
V169b	8.14	120.22	176.66	62.36	32.67
T170	7.61	122.90	175.02	63.33	70.85
A171	8.11	126.81	177.87	52.18	19.23
A171b	-	-	177.17	52.43	19.17
A172	8.08	125.02	-	50.53	17.91
A172b	7.86	123.04	-	50.48	19.39
P173	-	-	177.79	63.05	32.50

K174	8.22	122.30	176.72	56.47	33.02
K175	7.79	128.13	-	57.93	33.71
K175b	7.83	128.53	-	57.75	33.25

References

1. Powers, E. T. & Balch, W. E. Diversity in the origins of proteostasis networks-a driver for protein function in evolution. *Nat Rev Mol Cell Biol.* **14,** 237–248 (2013).

2. Balch, W. E., Morimoto, R. I., Dillin, A. & Kelly, J. W. Adapting Proteostatis for Disease Intervention. *Science (80-.).* **319,** 916–919 (2008).

3. Caspers, G. J., Leunissen, J. A. M. & de Jong, W. W. The expanding small heat-shock protein family, and structure predictions of the conserved "α-crystallin domain". *J. Mol. Evol.* **40,** 238–248 (1995).

4. van Montfort, R. L., Basha, E., Friedrich, K. L., Slingsby, C. & Vierling, E. Crystal structure and assembly of a eukaryotic small heat shock protein. *Nat. Struct. Biol.* **8,** 1025–1030 (2001).

5. Haslbeck, M., Franzmann, T., Weinfurtner, D. & Buchner, J. Some like it hot: The structure and function of small heat-shock proteins. *Nat. Struct. Mol. Biol.* **12,** 842–846 (2005).

6. de Jong, W. W., Caspers, G.-J. & Leunissen, J. A. M. Genealogy of the α-crystallin - small heat-shock protein superfamily. *Int. J. Biol. Macromol.* **22,** 151–162 (1998).

7. Jakob, U., Gaestel, M., Engel, K. & Buchner, J. Small Heat Shock Proteins Are Molecular Chaperones. *J. Biol. Chem.* **268,** 1517–1520 (1993).

8. Haslbeck, M. & Vierling, E. A First Line of Stress Defense: Small Heat Shock Proteins and Their Function in Protein Homeostasis. *J. Mol. Biol.* **427,** 1537–1548 (2015).

9. Bukau, B., Weissman, J. & Horwich, A. Molecular Chaperones and Protein Quality Control. *Cell* **125,** 443–451 (2006).

10. Ecroyd, H. & Carver, J. A. Crystallin proteins and amyloid fibrils. *Cell. Mol. Life Sci.* **66,** 62–81 (2009).

11. Vos, M. J. *et al.* HSPB7 is the most potent polyQ aggregation suppressor within the HSPB family of molecular chaperones. *Hum. Mol. Genet.* **19,** 4677–4693 (2010).

12. Ke, L. *et al.* HSPB1, HSPB6, HSPB7 and HSPB8 Protect against RhoA GTPase-

Induced Remodeling in Tachypaced Atrial Myocytes. *PLoS One* **6,** 1–10 (2011).

13. Narberhaus, F. α-Crystallin-Type Heat Shock Proteins: Socializing Minichaperones in the Context of a Multichaperone Network. *Microbiol. Mol. Biol. Rev.* **66,** 64–93 (2002).

14. Körner, C. T. Untersuchung der Proteinsubstanzen in den lichtbrechenden Medien des Auges. *Z. Physiol. Chem.* **18,** 61–106 (1894).

15. Siezen, R. J., Bindels, J. G. & Hoenders, H. J. The Quaternary Structure of Bovine α-Crystallin - Size and Charge Microheterogeneity: More than 1000 Different Hybrids? *Eur. J. Biochem.* **91,** 387–396 (1978).

16. Srinivas, P., Narahari, A., Petrash, J. M., Swamy, M. J. & Reddy, G. B. Importance of Eye Lens α-Crystallin Heteropolymer with 3:1 αA to αB Ratio: Stability, Aggregation, and Modifications. *IUBMB Life* **62,** 693–702 (2010).

17. Bloemendal, H. *et al.* Aging and vision: structure, stability and function of lens crystallin. *Prog. Biophys. Mol. Biol.* **86,** 407–485 (2004).

18. Clark, A. R., Lubsen, N. H. & Slingsby, C. sHSP in the eye lens: Crystallin mutations, cataract and proteostasis. *Int. J. Biochem. Cell Biol.* **44,** 1687–1697 (2012).

19. Plater, M. L., Goode, D. & Crabbe, M. J. C. Effects of Site-directed Mutations on the Chaperone-like Activity of αB-Crystallin. *J. Biol. Chem.* **271,** 28558–28566 (1996).

20. Beswick, H. T. & Harding, J. J. High-molecular-weight Crystallin Aggregate Formation Resulting from Non-Enzymic Carbamylation of Lens Crystallins: Relevance to Cataract Formation. *Exp. Eye Res.* **45,** 569–578 (1987).

21. Lowe, J. *et al.* αB crystallin expression in nonlenticular tissues and selective presence in ubiquitinated inclusion bodies in human disease. *J. Pathol.* **166,** 61–68 (1992).

22. Renkawek, K., Stege, G. J. J. & Bosman, G. J. C. G. M. Dementia, gliosis and expression of the small heat shock proteins hsp27 and αB-crystallin in Parkinson's disease. *Neuroreport* **10,** 2273–2276 (1999).

23. Cox, D. & Ecroyd, H. The small heat shock proteins αB-crystallin (HSPB5) and Hsp27 (HSPB1) inhibit the intracellular aggregation of α-synuclein. *Cell Stress Chaperones* **22,** 589–600 (2017).

24. Lowe, J. *et al.* Dementia with β-amyloid deposition involvement of αB-crystallin supports two main diseases. *Lancet* 515–516 (1990).

25. Cox, D., Carver, J. A. & Ecroyd, H. Preventing α-synuclein aggregation: The role of the small heat-shock molecular chaperone proteins. *Biochim. Biophys. Acta* **1842,** 1830–1843 (2014).

26. Shinohara, H., Inaguma, Y., Goto, S., Inagaki, T. & Kato, K. αB crystallin and HSP28 are enhanced in the cerebral cortex of patients with Alzheimer's disease. *J. Neurol. Sci.* **119,** 203–208 (1993).

27. Björkdahl, C. *et al.* Small heat shock proteins Hsp27 or αB-crystallin and the protein components of neurofibrillary tangles: Tau and neurofilaments. *J. Neurosci. Res.* **86,** 1343–1352 (2008).

28. Ito, H. *et al.* Phosphorylation-induced Change of the Oligomerization State of αB-crystallin. *J. Biol. Chem.* **276,** 5346–5352 (2001).

29. Kato, K. *et al.* Ser-59 is the major phosphorylation site in αB-crystallin accumulated in the brains of patients with Alexander's disease. *J. Neurochem.* **76,** 730–736 (2001).

30. Vicart, P. *et al.* A missense mutation in the αb-crystallin chaperone gene causes a desmin-related myopathy. *Nat. Genet.* **20,** 92–95 (1998).

31. Rajasekaran, N. S. *et al.* Human αB-Crystallin Mutation Causes Oxido-Reductive Stress and Protein Aggregation Cardiomyopathy in Mice. *Cell* **130,** 427–439 (2007).

32. Iwaki, T., Kume-lwaki, A., Liem, R. K. H. & Goldman, J. E. αB-Crystallin Is Expressed in Non-Lenticular Tissues and Accumulates in Alexander's Disease Brain. *Cell* **57,** 71–78 (1989).

33. Voduc, K. D. *et al.* αB-crystallin expression in breast cancer is associated with brain metastasis. *npj Breast Cancer* **1,** 15014 (2015).

34. Kim, M. S., Lee, H. W. & Lee, E. H. Renal tumor with alpha b crystallin expression. *Int. J. Clin. Exp. Pathol.* **8,** 9383–9389 (2015).

35. Shi, Q.-M. *et al.* High level of αB-crystallin contributes to the progression of osteosarcoma. *Oncotarget* **7,** 9007–9016 (2016).

36. Bennardini, F., Wrzosek, A. & Chiesi, M. αB-Crystallin in Cardiac Tissue: Association with actin and desmin Filaments. *Circ. Res.* **71,** 288–294 (1992).

37. Raman, B. *et al.* αB-crystallin, a small heat-shock protein, prevents the amyloid fibril growth of an amyloid β-peptide and β2-microglobulin. *Biochem. J.* **392,** 573–581 (2005).

38. Ousman, S. S. *et al.* Protective and therapeutic role for αB-crystallin in autoimmune demyelination. *Nat. Lett.* **448,** 474–481 (2007).

39. Iwaki, T. & Tateishi, J. Immunohistochemical Demonstration of AlphaB-Crystallin in Hamartomas of Tuberous Sclerosis. *Am. J. Pathol.* **139,** 1303–1308 (1991).

40. Lin, D. I. *et al.* Phosphorylation-Dependent Ubiquitination of Cyclin D1 by the SCFFBX4-αB Crystallin Complex. *Mol. Cell* **24,** 355–366 (2006).

41. Wignes, J. a, Goldman, J. W., Weihl, C. C., Bartley, M. G. & Andley, U. P. p62 expression and autophagy in αB-crystallin R120G mutant knock-in mouse model of hereditary cataract. *Exp. Eye Res.* **115,** 263–73 (2013).

42. Andley, U. P., Hamilton, P. D., Ravi, N. & Weihl, C. C. A Knock-In Mouse Model for the R120G Mutation of αB-Crystallin Recapitulates Human Hereditary Myopathy and Cataracts. *PLoS One* **6,** 1–13 (2011).

43. Liu, Y. *et al.* A Novel αB-Crystallin Mutation Associated with Autosomal Dominant Congenital Lamellar Cataract. *Invest. Ophthalmol. Vis. Sci.* **47,** 1069–1075 (2006).

44. Horwitz, J. Alpha-crystallin can function as a molecular chaperone. *PNAS* **89,** 10449–10453 (1992).

45. de Jong, W. W., Leunissen, J. A. M. & Voorter, C. E. M. Evolution of the α-Crystallin/Small Heat-Shock Protein Family. *Mol Biol Evol* **10,** 103–126 (1993).

46. Pasta, S. Y., Raman, B., Ramakrishna, T. & Rao, C. M. The IXI/V motif in the C-terminal extension of α-crystallins: alternative interactions and oligomeric assemblies. *Mol. Vis.* **10,** 655–662 (2004).

47. Duncan, M. K., Cvekl, A., Kantorow, M. & Pietigorsky, J. Lens crystallins, in F.J. Lovicu, M.L. Robinson (Eds.). Development of the Ocular Lens. *Cambridge Univ. Press. Cambridge* 119–150 (2004).

48. Hochberg, G. K. A. & Benesch, J. L. P. Dynamical structure of αB-crystallin. *Prog. Biophys. Mol. Biol.* **115,** 11–20 (2014).

49. Aquilina, J. A., Benesch, J. L. P., Bateman, O. A., Slingsby, C. & Robinson, C. V. Polydispersity of a mammalian chaperone: Mass spectrometry reveals the population of oligomers in αB-crystallin. *PNAS* **100,** 10611–10616 (2003).

50. Aquilina, J. A. *et al.* Phosphorylation of αB-crystallin Alters Chaperone Function through Loss of Dimeric Substructure. *J. Biol. Chem.* **279,** 28675–28680 (2004).

51. Benesch, J. L. P., Ayoub, M., Robinson, C. V. & Aquilina, J. A. Small Heat Shock Protein Activity Is Regulated by Variable Oligomeric Substructure. *J. Biol. Chem.* **283,** 28513–28517 (2008).

52. Spector, A., Freund, T., Li, L.-K. & Augusteyn, R. C. Age-dependent changes in the structure of alpha crystallin. *Investig. Ophtalmol.* **10,** 677–686 (1971).

53. Siezen, R. J., Bindels, J. G. & Hoelders, H. J. The Quaternary Structure of Bovine α-Crystallin. *Eur. J. Biochem.* **111,** 435–444 (1980).

54. Tardieu, A., Laporte, D., Licinio, P., Krop, B. & Delaye, M. Calf Lens α-Crystallin Quaternary Structure. *J. Mol. Biol.* **192,** 711–24 (1986).

55. Augusteyn, R. C. & Koretz, J. F. A possible structure for α-crystallin. *FEBS Lett.* **222,** 1–5 (1987).

56. Walsh, M. T., Sen, A. C. & Chakrabarti, B. Micellar Subunit Assembly in a Three-layer Model of Oligomeric α-Crystallin. *J. Biol. Chem.* **266,** 20079–20084 (1991).

57. Wistow, G. Possible Tetramer-based Quaternary Structure for α-Crystallins and Small Heat Shock Proteins. *Exp. Eye Res.* **56,** 729–732 (1993).

58. Carver, J. A., Aquilina, J. A. & Truscott, R. J. W. An investigation into the stability of α-crystallin by NMR spectroscopy; evidence for a two-domain structure. *Biochim. Biophys. Acta* **1164,** 22–28 (1993).

59. Carver, J. A., Aquilina, J. A., Truscott, R. J. W. & Ralston, G. B. Identification by 1H NMR spectroscopy of flexible C-terminal extensions in bovine lens α-crystallin. *FEBS Lett.* **311,** 143–149 (1992).

60. Carver, J. A., Aquilina, J. A. & Truscott, R. J. W. A Possible Chaperone-like Quaternary Structure for α-Crystallin. *Exp. Eye Res.* **59,** 231–234 (1994).

61. Haley, D. A., Horwitz, J. & Stewart, P. L. The Small Heat-shock Protein, αB-crystallin, has a Variable Quaternary Structure. *J. Mol. Biol.* **277,** 27–35

(1998).

62. Haley, D. A., Bova, M. P., Huang, Q. L., Mchaourab, H. S. & Stewart, P. L. Small Heat-Shock Protein Structures Reveal a Continuum from Symmetric to Variable Assemblies. *J. Mol. Biol.* **298,** 261–72 (2000).

63. Muchowski, P. J., Wu, G. J. S., Liang, J. J. N., Adman, E. T. & Clark, J. I. Site-directed Mutations Within the Core 'α-crystallin' Domain of the Small Heat-shock Protein, Human αB-crystallin, Decrease Molecular Chaperone Functions. *J. Mol. Biol.* **289,** 397–411 (1999).

64. Kim, K. K., Kim, R. & Kim, S. H. Crystal structure of a small heat-shock protein. *Lett. to Nat.* **394,** 595–599 (1998).

65. Koteiche, H. A. & Mchaourab, H. S. Folding Pattern of the α-Crystallin Domain in αA-Crystallin Determined by Site-Directed Spin Labeling. *J. Mol. Biol.* **294,** 561–577 (1999).

66. Feil, I. K., Malfois, M., Hendle, J., van Der Zandt, H. & Svergun, D. I. A Novel Quaternary Structure of the Dimeric α-crystallin Domain with Chaperone-like Activity. *J. Biol. Chem.* **276,** 12024–12029 (2001).

67. Jehle, S. *et al.* αB-Crystallin: A Hybrid Solid-State/Solution-State NMR Investigation Reveals Structural Aspects of the Heterogeneous Oligomer. *J. Mol. Biol.* **385,** 1481–1497 (2009).

68. Jehle, S. *et al.* Solid-state NMR and SAXS studies provide a structural basis for the activation of αB-crystallin oligomers. *Nat. Struct. Mol. Biol.* **17,** 1037–1042 (2010).

69. Bagnéris, C. *et al.* Crystal Structures of α-Crystallin Domain Dimers of αB-Crystallin and Hsp20. *J. Mol. Biol.* **392,** 1242–1252 (2009).

70. Laganowsky, A. *et al.* Crystal structures of truncated alphaA and alphaB crystallins reveal structural mechanisms of polydispersity important for eye lens function. *Protein Sci.* **19,** 1031–1043 (2010).

71. Hochberg, G. K. A. *et al.* The structured core domain of αB-crystallin can prevent amyloid fibrillation and associated toxicity. *Proc. Natl. Acad. Sci. USA* **111,** 1–9 (2014).

72. Rajagopal, P. *et al.* A conserved histidine modulates HSPB5 structure to trigger chaperone activity in response to stress related acidosis. *Elife* **4,** 1–21 (2015).

73. Baldwin, A. J., Lioe, H., Robinson, C. V, Kay, L. E. & Benesch, J. L. P. αB-crystallin polydispersity is a consequence of unbiased quaternary dynamics. *J. Mol. Biol.* **413,** 297–309 (2011).

74. Mainz, A. *et al.* The chaperone αB-crystallin uses different interfaces to capture an amorphous and an amyloid client. *Nat. Struct. Mol. Biol.* **22,** 898–907 (2015).

75. Peschek, J. *et al.* The eye lens chaperone α-crystallin forms defined globular assemblies. *PNAS* **106,** 13272–13277 (2009).

76. Jehle, S. *et al.* N-terminal domain of αB-crystallin provides a conformational switch for multimerization and structural heterogeneity. *PNAS* **108,** 6409–6414 (2011).

77. Braun, N. *et al.* Multiple molecular architectures of the eye lens chaperone αB-crystallin elucidated by a triple hybrid approach. *PNAS* **108,** 20491–20496 (2011).

78. Baldwin, A. J. *et al.* Probing Dynamic Conformations of the High-Molecular-Weight αB-Crystallin Heat Shock Protein Ensemble by NMR Spectroscopy. *JACS* **134,** 15343–15350 (2012).

79. Ingolia, T. D. & Craig, E. A. Four small Drosophila heat shock proteins are related to each other and to mammalian α-crystallin. *PNAS* **79,** 2360–2364 (1982).

80. Wistow, G. Domain structure and evolution in α-crystallins and small heat-shock proteins. *FEBS Lett.* **181,** 1–6 (1985).

81. Poulain, P., Gelly, J. C. & Flatters, D. Detection and Architecture of Small Heat Shock Protein Monomers. *PLoS One* **5,** 1–10 (2010).

82. Kappé, G. *et al.* The human genome encodes 10 α-crystallin-related small heat shock proteins: HspB1-10. *Cell Stress Chaperones* **8,** 53–61 (2003).

83. Fu, L. & Liang, J. J.-N. Enhanced stability of αB-crystallin in the presence of small heat shock protein Hsp27. *Biochem. Biophys. Res. Commun.* **302,** 710–714 (2003).

84. Mymrikov, E. V., Seit-Nebi, A. S. & Gusev, N. B. Heterooligomeric complexes of human small heat shock proteins. *Cell Stress Chaperones* **17,** 157–169 (2012).

85.   Fontaine, J.-M., Sun, X., Benndorf, R. & Welsh, M. J. Interactions of HSP22 (HSPB8) with HSP20, αB-crystallin, and HSPB3. *Biochem. Biophys. Res. Commun.* **337,** 1006–1011 (2005).

86.   Liu, Z. *et al.* Mechanistic insights into the switch of αB-crystallin chaperone activity and self-multimerization. *J. Biol. Chem.* **293,** 14880–14890 (2018).

87.   Tyedmers, J., Mogk, A. & Bukau, B. Cellular strategies for controlling protein aggregation. *Nat. Rev. Mol. Cell Biol.* **11,** 777–788 (2010).

88.   Treweek, T. M., Meehan, S., Ecroyd, H. & Carver, J. A. Small heat-shock proteins: Important players in regulating cellular proteostasis. *Cell. Mol. Life Sci.* **72,** 429–451 (2015).

89.   Abgar, S., Vanhoudt, J., Aerts, T. & Clauwaert, J. Study of the Chaperoning Mechanism of Bovine Lens α-Crystallin, a Member of the α-Small Heat Shock Superfamily. *Biophys. J.* **80,** 1986–1995 (2001).

90.   Reddy, G. B., Reddy, P. Y. & Suryanarayana, P. αA- and αB-Crystallins Protect Glucose-6-Phosphate Dehydrogenase against UVB Irradiation-Induced Inactivation. *Biochem. Biophys. Res. Commun.* **282,** 712–716 (2001).

91.   Raman, B. & Rao, C. M. Chaperone-like Activity and Quaternary Structure of α-Crystallin. *J. Biol. Chem.* **269,** 27264–27268 (1994).

92.   Yaung, J. *et al.* α-Crystallin distribution in retinal pigment epithelium and effect of gene knockouts on sensitivity to oxidative stress. *Mol. Vis.* **13,** 566–577 (2007).

93.   Ehrnsperger, M., Gräber, S., Gaestel, M. & Buchner, J. Binding of non-native protein to Hsp25 during heat shock creates a reservoir of folding intermediates for reactivation. *EMBO J.* **16,** 221–229 (1997).

94.   Horwitz, J. The Function of Alpha-Crystallin. *Proctor Lect.* **34,** 10–22 (1993).

95.   Jovcevski, B., Aquilina, J. A., Benesch, J. L. P. & Ecroyd, H. The influence of the N-terminal region proximal to the core domain on the assembly and chaperone activity of αB-crystallin. *Cell Stress Chaperones* **23,** 827–836 (2018).

96.   Cox, D., Selig, E., Griffin, M. D. W., Carver, J. A. & Ecroyd, H. Small Heat-shock Proteins Prevent α-Synuclein Aggregation via Transient Interactions and Their Efficacy Is Affected by the Rate of Aggregation. *J. Biol. Chem.* **291,** 22618–22629 (2016).

97. Singh, B. N., Rao, K. S., Ramakrishna, T., Rangaraj, N. & Rao, C. M. Association of αB-Crystallin, a Small Heat Shock Protein, with Actin: Role in Modulating Actin Filament Dynamics in Vivo. *J. Mol. Biol.* **366,** 756–767 (2007).

98. Bullard, B. *et al.* Association of the Chaperone αB-crystallin with Titin in Heart Muscle. *J. Biol. Chem.* **279,** 7917–7924 (2004).

99. Merck, K. B., de Haard-Hoekman, W. A., Oude Essink, B. B., Bloemendal, H. & de Jong, W. W. Expression and aggregation of recombinant αA-crystallin and its two domains. *Biochim. Biophys. Acta* **1130,** 267–276 (1992).

100. Aquilina, J. A. & Watt, S. J. The N-terminal domain of αB-crystallin is protected from proteolysis by bound substrate. *Biochem. Biophys. Res. Commun.* **353,** 1115–1120 (2007).

101. Ghosh, J. G., Shenoy, A. K. & Clark, J. I. N- and C-terminal motifs in Human αB Crystallin Play an Important Role in the Recognition, Selection, and Solubilization of Substrates. *Biochemistry* **45,** 13847–13854 (2006).

102. Ghosh, J. G., Estrada, M. R. & Clark, J. I. Interactive Domains for Chaperone Activity in the Small Heat Shock Protein, Human αB crystallin. *Biochemistry* **44,** 14854–14869 (2005).

103. Ghosh, J. G., Estrada, M. R., Houck, S. A. & Clark, J. I. The function of the β3 interactive domain in the small heat shock protein and molecular chaperone, human αB crystallin. *Cell Stress Chaperones* **11,** 187–197 (2006).

104. Treweek, T. M. *et al.* Site-Directed Mutations in the C-terminal Extension of Human αB-Crystallin Affect Chaperone Function and Block Amyloid Fibril Formation. *PLoS One* **10,** 1–10 (2007).

105. Giese, K. C. & Vierling, E. Mutants in a small heat shock protein that affect the oligomeric state: Analysis and allele-specific suppression. *J. Biol. Chem.* **279,** 32674–32683 (2004).

106. Haslbeck, M. Hsp26: a temperature-regulated chaperone. *EMBO J.* **18,** 6744–6751 (1999).

107. Shashidharamurthy, R., Koteiche, H. A., Dong, J. & Mchaourab, H. S. Mechanism of chaperone function in small heat shock proteins: Dissociation of the HSP27 oligomer is required for recognition and binding of destabilized T4 lysozyme. *J. Biol. Chem.* **280,** 5281–5289 (2005).

108. Miesbauer, L. R. *et al.* Post-translational Modifications of Water-Soluble Human Lens Crystallins from Young Adults. *J. Biol. Chem.* **269,** 12494–12502 (1994).

109. Ecroyd, H. *et al.* Mimicking phosphorylation of αB-crystallin affects its chaperone activity. *Biochem. J.* **401,** 129–141 (2007).

110. Peschek, J. *et al.* Regulated structural transitions unleash the chaperone activity of αB-crystallin. *PNAS* **110,** E3780–E37889 (2013).

111. Mchaourab, H. S., Godar, J. A. & Stewart, P. L. Structure and mechanism of protein stability sensors: Chaperone activity of small heat shock proteins. *Biochemistry* **48,** 3828–3837 (2009).

112. Wang, K., Gawinowicz, M. A. & Spector, A. The Effect of Stress on the Pattern of Phosphorylation of αA and αB crystallin in the Rat Lens. *Exp. Eye Res.* **71,** 385–393 (2000).

113. Nicholl, I. D. & Quinlan, R. A. Chaperone activity of α-crystallins modulates intermediate filament assembly. *EMBO J.* **13,** 945–953 (1994).

114. Wang, K. & Spector, A. α-Crystallin stabilizes actin filaments and prevents cytochalasin-induced depolymerization in a phosphorylation-dependent manner. *Eur. J. Biochem.* **242,** 56–66 (1996).

115. Ahmad, F., Raman, B., Ramakrishna, T. & Rao, C. M. Effect of Phosphorylation on αB-crystallin: Differences in Stability, Subunit Exchange and Chaperone Activity of Homo and Mixed Oligomers of αB-Crystallin and its Phosphorylation-mimicking Mutant. *J. Mol. Biol.* **375,** 1040–1051 (2008).

116. Bakthisaran, R., Akula, K. K., Tangirala, R. & Rao, C. M. Phosphorylation of αB-crystallin: Role in stress, aging and patho-physiological conditions. *Biochim. Biophys. Acta - Gen. Subj.* **1860,** 167–182 (2016).

117. Muranova, L. K., Sudnitsyna, M. V. & Gusev, N. B. αB-Crystallin Phosphorylation: Advances and Problems. *Biochem.* **83,** 1196–1206 (2018).

118. Arrigo, A.-P. Human small heat shock proteins: Protein interactomes of homo- and hetero-oligomeric complexes: An update. *FEBS Lett.* **587,** 1959–1969 (2013).

119. Chebotareva, N. A. *et al.* Oligomeric state of αB-crystallin under crowded conditions. *Biochem. Biophys. Res. Commun.* **508,** 1101–1105 (2018).

120. Levitt, M. H. *Spin Dynamics: Basics of Nuclear Magnetic Resonance.* (John Wiley & Sons, 2013).

121. Duer, M. J. *Solid State NMR Spetroscopy: Principles and Applications.* (John Wiley & Sons, 2008).

122. Keeler, J. *Understanding NMR Spectroscopy.* (John Wiley & Sons, 2011).

123. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Skelton, N. J. & Rance, M. *Protein NMR Spectroscopy: Principles and Practice.* (Elsevier, 2010).

124. Ernst, R. R., Bodenhausen, G. & Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimension.* (Oxford University Press, 1990).

125. Zeeman, P. The effect of magnetisation on the nature of light emitted by a substance. *Nature* **55,** 347 (1897).

126. Bloch, F., Hansen, W. W. & Packard, M. Nuclear Induction. *Phys. Rev.* **69,** 127 (1946).

127. Bloch, F. Nuclear Induction. *Phys. Rev.* **70,** 460–474 (1946).

128. Purcell, E. M., Torrey, H. C. & Pound, R. V. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys. Rev.* **69,** 37–38 (1946).

129. Hore, P. J. *Nuclear Magnetic Resonance.* (Oxford University Press, 1989).

130. Ulrich, E. L. *et al.* BioMagResBank. *Nucleic Acids Res.* **36,** D402–D408 (2007).

131. Redfield, C. Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. *Methods* **34,** 121–132 (2004).

132. Ernst, R. R. & Anderson, W. A. Application of Fourier Transform Spectroscopy to Magnetic Resonance. *Rev. Sci. Instrum.* **37,** 93–102 (1966).

133. Andrew, E. R., Bradbury, A. & Eades, R. G. Nuclear Magnetic Resonance Spectra from a Crystal rotated at High Speed. *Nature* **162,** 1659 (1958).

134. Baldus, M. Solid-State NMR Spectroscopy: Molecular Structure and Organization at the Atomic Level. *Angew. Chemie Int. Ed.* **45,** 1186–1188 (2006).

135. Wickramasinghe, A. *et al.* Evolution of CPMAS under fast magic-angle-spinning at 100 kHz and beyond. *Solid State Nucl. Magn. Reson.* **72,** 9–16 (2015).

136. Struppe, J. *et al.* Expanding the Horizons for Structural Analysis of Fully Protonated Protein Assemblies by NMR Spectroscopy at MAS Frequencies Above 100 kHz. *Solid State Nucl. Magn. Reson.* **87,** 117–125 (2017).

137. Pines, A., Gibby, M. G. & Waugh, J. S. Proton-enhanced NMR of dilute spins in solids. *J. Chem. Phys.* **569,** 569–590 (1973).

138. Hartmann, S. & Hahn, E. Nuclear Double Resonance in the Rotating Frame. *Phys. Rev.* **128,** 2042–2053 (1962).

139. Metz, G., Wu, X. L. & Smith, S. O. Ramped Amplitude Cross Polarization in Magic Angle Spinning NMR. *J. Magn. Reson. Ser. A* **110,** 219–227 (1994).

140. Szeverenyi, N. M., Sullivan, M. J. & Maciel, G. E. Observation of spin exchange by two-dimensional fourier transform 13C cross polarization-magic-angle spinning. *J. Magn. Reson.* **47,** 462–475 (1982).

141. Bloembergen, N. On the Interaction of Nuclear Spins in a Crystalline Lattice. *Physica* **XV,** 386–426 (1949).

142. Bénédicte, E., Lesage, A., Steuernagel, S., Böckmann, A. & Emsley, L. Proton to Carbon-13 INEPT in Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **127,** 17296–17302 (2005).

143. Baldus, M., Geurts, D. G., Hediger, S. & Meier, B. H. Efficient 15N-13C Polarization Transfer by Adiabatic-Passage Hartmann-Hahn Cross Polarization. *J. Magn. Reson. Ser. A1* **118,** 140–144 (1996).

144. Castellani, F. *et al.* Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **420,** 98–102 (2002).

145. Loquet, A., Lv, G., Giller, K., Becker, S. & Lange, A. 13C spin dilution for simplified and complete solid-state NMR resonance assignment of insoluble biological assemblies. *JACS* **133,** 4722–4725 (2011).

146. Hong, M. & Jakes, K. Selective and extensive 13C labeling of a membrane protein for solid-state NMR investigations. 71–74 (1999).

147. Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45,** 342–350 (2012).

148. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25,** 495–503 (1992).

149. Jeschke, G. *et al.* DeerAnalysis2006 - a Comprehensive Software Packacge for Analyzing Pulsed ELDOR Data. *Appl. Magn. Reson.* **30,** 473–498 (2006).

150. Fontaine, J.-M., Rest, J. S., Welsh, M. J. & Benndorf, R. The sperm outer dense

fiber protein is the 10th member of the superfamily of mammalian small stress proteins. *Cell Stress Chaperones* **8,** 62–69 (2003).

151. Stamler, R., Kappé, G., Boelens, W. & Slingsby, C. Wrapping the α-Crystallin Domain Fold in a Chaperone Assembly. *J. Mol. Biol.* **353,** 68–79 (2005).

152. Aquilina, J. A. *et al.* Subunit Exchange of Polydisperse Proteins. *J. Biol. Chem.* **280,** 14485–14491 (2005).

153. Sengupta, I., Nadaud, P. S. & Jaroniec, C. P. Protein Structure Determination with Paramagnetic Solid-State NMR Spectroscopy. *Acc. Chem. Res.* **46,** 2117–2126 (2013).

154. Bertini, I., Luchinat, C. & Parigi, G. *Solution NMR of Paramagnetic Molecules*. (2001).

155. Pintacuda, G. & Kervern, G. Paramagnetic Solid-State Magic-Angle Spinning NMR Spectroscopy. in *Modern NMR Methodology* 157–200 (2012).

156. Otting, G. Protein NMR Using Paramagnetic Ions. *Annu. Rev. Biophys.* **39,** 387–405 (2010).

157. Banci, L., Bertini, I., Luchinat, C. & Mori, M. NMR in structural proteomics and beyond. *Prog. Nucl. Magn. Reson. Spectrosc.* **56,** 247–266 (2010).

158. Nagaraj, M. *et al.* Surface Binding of TOTAPOL Assists Structural Investigations of Amyloid Fibrils by Dynamic Nuclear Polarization NMR Spectroscopy. *ChemBioChem* **17,** 1308–1311 (2016).

159. Rogawski, R. *et al.* NMR Signal Quenching from Bound Biradical Affinity Reagents in DNP Samples. *J. Phys. Chem. B* **121,** 10770–10781 (2017).

160. Ravera, E. *et al.* DNP-Enhanced MAS NMR of Bovine Serum Albumin Sediments and Solutions. *J. Phys. Chem. B* **118,** 2957–2965 (2014).

161. Thampi, P. & Abraham, E. C. Influence of the C-terminal Residues on Oligomerization of αA-crystallin. *Biochemistry* **42,** 11857–11863 (2003).

162. Baldwin, A. J. *et al.* Quaternary Dynamics of αB-Crystallin as a Direct Consequence of Localised Tertiary Fluctuations in the C-Terminus. *J. Mol. Biol.* **413,** 310–320 (2011).

163. Alderson, T. R., Benesch, J. L. P. & Baldwin, A. J. Proline isomerization in the C-terminal region of HSP27. *Cell Stress Chaperones* **22,** 639–651 (2017).

164. Murugesan, R., Santhoshkumar, P. & Sharma, K. K. Role of αBI5 and αBT162

residues in subunit interaction during oligomerization of αB-crystallin. *Mol. Vis.* **14,** 1835–1844 (2008).

165. Heirbaut, M. *et al.* Specific sequences in the N-terminal domain of human small heat-shock protein HSPB6 dictate preferential hetero-oligomerization with the orthologue HSPB1. *J. Biol. Chem.* **292,** 9944–9957 (2017).

166. Arntson, K. E. & Pomerantz, W. C. K. Protein-Observed Fluorine NMR: A Bioorthogonal Approach for Small Molecule Discovery. *J. Med. Chem.* **59,** 5158–5171 (2016).

List of figures

List of tables

Summary

The ubiquitous αB-crystallin belongs to the small heat shock proteins, a protein class involved in the maintenance of the homeostasis in cells. As a chaperone it prevents the aggregation of misfolded proteins by keeping them in the dissolved state. αB-crystallin is found upregulated among others in neurodegenerative disease states like Alzheimer's and Parkinson and cardiomyopathies as well as certain cancer types.

The 175 amino-acid chaperone forms spherical oligomers with particle size ranging from about 10 to 40 subunits, with the maximum being found at 26-28 protomers per particle. The sequence can be divided into three domains, the ~60 residue N-terminal domain, the α-crystallin domain, and a 25 amino acid C-terminal domain with a highly conserved IXI-motif. The hydrophilicity of the CTD provides the protein complex its excellent solubility properties. All three domains of the protein contribute to the formation of higher order oligomers, as short constructs do not lead to oligomers with the same particle distribution as the wild-type form of αB-crystallin. The α-crystallin domain, the common feature to all small heat shock proteins, is structurally highly conserved and comprises six β-sheets, β3 to β9, arranged in an immunoglobulin-like fold. Two crystallin domains form a dimer, via an antiparallel arrangement of their long β6+7 strands, which can occur in three different registers. This interface is described as the lowest level of building blocks for the oligomerization and is very pH-sensitive. Binding of the C-terminal IXI-motif to a hydrophobic groove formed by β4 and β8-strands in the core is considered the second building element in oligomer formation. Respective short constructs comprising the C-terminal motif and the core domain can form hexamers. Heterogenous interactions mediated through the hydrophobic N-termini lead to the complex oligomerization behavior of the wild-type protein. However, the exact way how the N-terminus contributes to the oligomer-formation is rather enigmatic. Furthermore, the degree of binding of the C-terminal IXI-motif to the hydrophobic groove is subject of controversy in the research on αB-crystallin.

Two models for a tetrahedral 24mer were developed, based on solid-state NMR results on the crystallin domain, cryo-EM and modelling as well as sequence

alignment. The models differ to a certain extend in the packing of the core-domain dimers as well as the positioning of the N- and C-termini.

As shown in this work, the IXI-motif is a conserved feature not only for the C-terminus but for the N-terminus as well, as it is present in five out of ten human small heat shock proteins. Through investigations on WT protein and a set of valine-mutants of the two IXI-motifs, namely αB I3/5V, αB I159/161V and αB I3/5/159/161V via MAS solid-state and solution NMR-spectroscopy methods, it was possible to identify and assign the N-terminal motif-related isoleucines for the first time. Carbon correlation spectra revealed very similar chemical shifts for the isoleucine sidechain atoms for both motifs, which is interpreted as both motifs being located in a very similar chemical environment within the complex oligomer. This strongly implicates that the N-terminal IXI-motif can also bind to the hydrophobic groove. Furthermore, through carbon direct-excitation experiments in solution NMR and subsequent subtraction of the resulting spectra of the different samples, the amount of bound C-terminal motif could be determined via integration of the signals, as well as the ratio of bound N- and C-termini. We found that the C-terminus is being bound to an extend of approximately 50% in the temperature range of 290 – 310 K. The binding of the N-terminus is more temperature dependent, the lower the temperature, the more of the N-terminal IXI-motif is bound to the groove. The binding of the N-terminus provides an additional way for self-regulation and oligomerization. This provides a rational explanation for the broad distribution of oligomers.

To probe the proximity of the IXI-motifs to the α-crystallin domain we mutated the residue between the isoleucines to cysteines (αB A4C and αB P160C). These mutants were spin-labelled with MTSL and in case of the C-terminal variant also with a TOTAPOL derivative. It turned out that the label is cleaved off very quickly in case of the N-terminal mutant, implying that the cysteines are very close to each other. In case of the C-terminal spin labelled variants, the paramagnetic relaxation enhancement-effect on the signals in the PDSD spectra was analyzed, the effect evaluated and discussed on the two existing 24mer models (Jehle, Weinkauf).

Zusammenfassung

Das ubiquitäre αB-Crystallin gehört zu der Familie der Hitzeschockproteine, eine Protein-Klasse die beteiligt ist, das Gleichgewicht der Proteine in der Zelle aufrecht zu erhalten. Als Chaperon verhindert es die Aggregation von fehlgefalteten Proteinen durch Stabilisierung der gelösten Form. αB-Crystallin ist unter anderem in Neurodegenerativen Erkrankungen sowie bestimmten Krebsarten hochreguliert

Das 175 Aminosäuren Chaperon bildet sphärische Partikel aus etwa 10-40 Untereinheiten, mit Maxima im Bereich von 26-28 Protomeren pro Partikel. Die Sequenz kann in drei Teile eingeteilt werden, den ~60 Reste lange N-terminale Domäne, die α-Crystallin-Domäne und eine 25 Reste lange C-terminale Domäne mit einem hoch konservierten IXI-Motiv. Die Hydrophilie der CTD verleiht dem Protein seine ausgezeichneten Lösungseigenschaften. Alle drei Domänen sind an der Bildung von höhergeordneten Oligomeren beteiligt, denn kurze Konstrukte führen nicht zu Partikeln mit gleicher Größenverteilung wie beim Wild-Typ von αB-Crystallin. Die α-Crystallin-Domäne, eine Gemeinsamkeit aller Hitzeschockproteine, ist strukturell hoch konserviert und beinhaltet sechs β-Faltblätter, β3 bis β9, aufgebaut in einer Immunoglobulin-ähnlichen Faltung. Zwei Crystallin-Domänen bilden ein Dimer durch die antiparallele Anordnung ihrer langen β6+7 Faltblätter, welche in drei verschiedenen Registern stattfinden kann. Diese Kontaktfläche ist als erster Baustein für die Oligomerisierung beschrieben und sehr pH-sensitiv. Die Bindung des C-terminalen IXI-Motives in einer hydrophoben Bindetasche geformt durch β4 und β8 wird als zweiter Baustein für die Oligomerisierung beschrieben. Entsprechende verkürzte Konstrukte, welche das C-terminale Motiv und die Crystallin-Domäne enthalten, können Hexamere bilden. Heterogene Interaktionen, vermittelt durch die N-Termini, führen zum komplexen Oligomerisierungs-Verhalten des Wild-Typ Proteins. Nichtsdestotrotz ist der exakte Beitrag des N-Terminus zur Oligomerisierung rätselhaft. Darüber hinaus ist der Anteil der Bindung des C-terminalen IXI-Motives in der hydrophoben Bindetasche kontrovers diskutiert.

Es wurden zwei Modelle für ein tetrahedrales 24mer entwickelt, basierend auf Festkörper-NMR Ergebnissen für die Crystallin-Domäne, Cryo-EM und Modellierung sowie Vergleichen von Sequenzen. Die Modelle unterscheiden sich zu einem

gewissen Grad in der Packung der Hauptdomäne sowie der Positionierung der N- und C-Termini.

Wie in dieser Arbeit gezeigt wird, ist das IXI-Motiv nicht im C-Terminus hoch konserviert, sondern auch im N-Terminus, denn es ist in fünf von zehn humanen Hitzeschockproteinen vorhanden. Durch Untersuchungen des Wild-Typs und einigen Valin-Mutanten der IXI-Motive, nämlich αB I3/5V, αB I159/161V und αB I3/5/159/161V, mittels MAS Festkörper-NMR und Lösungs-NMR-Spektroskopie, war es möglich, die N-terminalen Isoleucine des IXI-Motives erstmals zu identifizieren und zuzuordnen. Kohlenstoff-Korrelations-Spektren offenbaren sehr ähnliche chemische Verschiebungen für die Seitenketten-Atome beider Motive, was dahingehend interpretiert wird, dass sich beide Motive in ähnlichen chemischen Umgebungen im komplexen Oligomer befinden. Dies ist ein starkes Indiz dafür, dass das N-terminale IXI-Motiv ebenfalls in die hydrophobe Bindetasche bindet. Zudem konnte durch Kohlenstoff-Direkt-Anregung und Subtraktion der erhaltenen Spektren für die verschiedenen Proben der Anteil an gebundenem C-terminalen Motiv durch Integration sowie das Verhältnis von gebundenem N- und C-Termini bestimmt werden. Wir haben herausgefunden, dass der C-Terminus zu einem Anteil von etwa 50% im Temperatur-Bereich von 290 – 310 K gebunden ist. Die Bindung des N-Terminus ist mehr temperaturabhängig, je niedriger die Temperatur, desto mehr N-terminales IXI-Motiv ist gebunden. Die Bindung des N-terminalen IXI-Motives birgt einen weiteren Baustein der Selbst-Regulation und Oligomerisierung. Dieses Ergebnis liefert eine rationale Erklärung für die breite Oligomer-Verteilung.

Um die Nähe der IXI-Motive zur α-Crystallin-Domäne zu testen, haben wir die Reste zwischen den Isoleucinen der Motive zu Cysteinen mutiert (αB A4C und αB P160C). Diese Mutanten wurden Spin-markiert mit MTSL und im Falle der C-terminalen Variante ebenfalls mit einem TOTAPOL-derivat. Es hat sich herausgestellt, dass im Falle der N-terminalen Mutante das paramagnetische Label sehr schnell abgespalten wird, was bedeutet, dass die Cysteine in räumlicher Nähe zueinander sind. Im Falle der C-terminalen Spin-Label-Varianten wurde der paramagnetische Ralaxations-Beschleunigungs-Effekt in PDSD Spektren analysiert, der Effekt evaluiert und an den beiden existierenden 24mer-Modellen (Jehle, Weinkauf) diskutiert.

Acknowledgements

List of publications

Not related to this work:

'Investigations on the mode of action of geohyronic acid, an inhibitor of eukaryotic protein translation from myxobacteria' Muthukumar, Y., **Münkemer, L. J.**, Mathieu, D., Richter, C., Schwalbe, H., Steinmetz, H., Kessler, W., Reichelt, J., Beutling, U., Frank, R., Büssow, K., van den Heuvel, J., Brönstrup, M., Tayler, R. E., Laschat, S., Sasse, F., PLoS One, 2018 July 31, 13(7): e0201605.

Curriculum Vitae

For reasons of data protection, the Curriculum Vitae is not published in the online version.

Declaration of honesty / Eidesstattliche Erklärung

Hereby I declare that I have written this thesis by myself, marked the sources of any quotations or content obtained otherwise and mentioned any personal help by name.

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig verfasst, die wörtlich oder inhaltlich anderen Quellen entnommenen Stellen als solche kenntlich gemacht und die Inanspruchnahme persönlicher Hilfe namentlich aufgeführt habe.

Ort, Datum   Lea Johanna Münkemer