Technische Universität München
Institut für Organische Chemie und Biochemie

Max-Planck-Institut für Biochemie
Abteilung Strukturforschung

Nuclear Magnetic Resonance Spectroscopic Investigations on the Green Fluorescent Protein, the Cyclase Associated Protein and Proteins Involved in Cancer Developement

Till Rehm

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. St. J. Glaser

Prüfer der Dissertation:

1. Univ.-Prof. Dr. H. Kessler
2. apl. Prof. Dr. L. Moroder

Die Dissertation wurde am 30.10.2002 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 21.11.2002 angenommen.

Publications

Parts of this thesis have already been published or will be published in due course:

Rehm, T., Huber R. and Holak T.A. *Application of NMR in structural proteomics: Screening for proteins amenable to structural analysis*, Structure, 10:1613-1618, 2002.

Stoll R., Renner C., Hansen S., Palme S., Klein C., Belling A., Zeslawski W., Kamionka M., *Rehm T.*, Mühlhahn P., Schumacher R., Hesse F., Kaluza B., Voelter W., Engh R.A. and Holak T.A. *Chalcone derivatives antagonize interactions between the human oncoprotein MDM2 and p53.*, Biochemistry, 40(2):336-344, 2001.

Kamionka, M., *Rehm, T.*, Beisel H.G., Lang K., Engh R.A. and Holak T.A. *In silico and NMR Identification of inhibitors of the IGF-I and IGF-Binding Protein 5 interaction*, J. Med. Chem., 45(26):5655-5660, 2002.

Brüggert, M., *Rehm, T.*, Shanker, S. and Holak, T.A. *A novel medium for expression of selectively ^{15}N labled proteins in SF9 insect cells*, J. Biomol. NMR, in press 2003.

Rehm, T., Mavoungou, C. Israel, L., Schleicher, M. and Holak T.A.: *Letter to the Editor: Sequence-specific (^{1}H, ^{15}N, ^{13}C) resonance assignment of the N-terminal domain of the cyclase-associated protein (CAP) from Dictyostelium discoideum*, J. Biomol. NMR, 23:337-338, 2002.

Rehm, T., Mavoungou, C. Israel, L., Ksiazek, D., Schleicher, M. and Holak T.A.: *Solution structure of the adenylyl cyclase associated protein (CAP) from Dictyostelium discoideum*, in preparation

Seifert, M.H.J., Georgescu, J., Ksiazek, D., Smialowski, P., *Rehm, T.*, Steipe, B. and Holak T.A. *Backbone Dynamics of Green Fluorescent Protein and the Effect of Histidine 148 Substitution*, Biochemistry, in press, 2003.

Georgescu, J., *Rehm, T.*, Wiehler, J., Steipe, B. and Holak T.A. *Letter to the Editor: Backbone H^N, N, C^α, and C^β assignment of the GFPuv mutant*, J. Biomol. NMR, in press, 2003.

Abbreviations

Amino acids are abbreviated according to either one or three letter IUPAC (International Union of Pure and Applied Chemistry) code. The customary acronyms are used for the NMR experiments.

2D	two-dimensional
3D	three-dimensional
CAP	cyclase associated protein
DMSO	dimethylsulfoxide
DMSO-d6	deuterated dimethylsulfoxide
DNA	deoxyribonucleic acid
ELISA	enzyme-linked immunosorbant assay
FMOC	9-Fluorenylmethoxycarbonyl
GFP	green flourescent protein
GST	glutathione-S-transferase protein
HSQC	heteronuclear single-quantum coherence
IC_{50}	inhibitior concentration wit 50% inhibition
IGF-I	insulin-like growth factor-I
IGFBP-4	IGF binding protein-4
IGFBP-5	IGF-binding protein-5
IPTG	Isopropylthiogalactosid
K_D	Dissociation constant
kDa	Kilodalton
MDM2	human murine double minute clone 2 protein

Ni-NTA	Ni-Nitrilotriacetic acid
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser enhancement
OD_{600}	optical density at 600 nm
PBS	phosphate buffered saline
PDB	Protein Data Bank
PIP_2	phosphatidylinositol 4,5-bisphosphate
ppm	parts per million
Sf9	spodoptera frugiperda
SH3	Src homology 3 domain
Tris	C,C,C-Tris(hydroxymethyl)-aminomethan

Contents

Chapter 1

Introduction

NMR Spectroscopy on Proteins in solution

Multi-dimensional nuclear magnetic resonance spectroscopy (NMR, for a list of abbriviations see appendix) is widely used for determining three-dimensional structures of small and medium sized proteins in solution. As NMR observes signals from individual atoms in these complex macromolecules, it is also possible to investigate further properties with atomic resolution. Among these are the mapping of binding sites of ligands, monitoring of folding and aggregation, detection of multiple conformations and dynamical properties. In this thesis no general description of the theory and application of the NMR method will be given as it was felt, that a comprehensive introduction into the wide field of high resolution, multidimensional NMR is outside the scope of this work and also because there are excellent books, covering all aspects of modern NMR techniques. The following text books are highly recommended for understanding the power of NMR:

Canet (1996) *Nuclear Magnetic Resonance. Concepts and Methods*; John Wiley & Sons, New York
[a nice introduction]

Croasmun & Carlson (1994) *Two dimensional NMR Spectroscopy. Applications for Chemists and Biochemists*, VCH Publisher, Weinheim
[an extensive treatise on multidimensional experiments]

Derome (1987) *Modern NMR Techniques for Chemistry Research*, Pergamon Press, Oxford
[a more technical approach]

Ernst, Bodenhausen & Wokaun (1997) *Principles of Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford
[from the *father* of NMR]

Evans (1995) *Biomolecular NMR Spectroscopy*, Oxford Univ. Press
[the biomolecular aspects of NMR]

Freeman (1988) *A Handbook of Nuclear Magnetic Resonance*, Longman, Essex
[an encyclopedia like reference book]

James & Oppenheimer (1994) *Nuclear Magnetic Resonance*, Academic Press, New York

James, Dötsch & Schmitz (2001) *Nuclear Magnetic Resonance of Biological Macromolecules, Part A and Part B*, Academic Press, New York
[three volumes on NMR of the excellent monography series *Methods in Enzymology*]

Neuhaus & Williamson (2000) *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH Publisher, New York
[a comprehensive study of the NOE Effect]

Cavanagh, Fairbrother, Palmer III & Skelton (1996) *Protein NMR Spectroscopy. Principles and Practice*, Academic Press, New York
[one of the best]

Reid (1997) *Protein NMR Techniques*, Humana Press, Totowa
[a very practical approach]

Wüthrich (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York
[the *Bible* of Protein NMR]

Scope of this work

This thesis is a collection of several NMR projects conducted at the Department of Structural Research of the Max Planck Institute of Biochemistry in the recent years. The wide range of applications of NMR in biochemical research is reflected in the diversity of the used methods. In Chapter 2, a general overview of the role NMR spectroscopy plays in structural proteomics is given. Especially the advantages of using NMR as a screening tool for proteins that can be

subjected to structural characterization by both NMR spectroscopy and X-ray crystallography are reviewed.

NMR is not only an excellent tool to select proteins amenable to structural analysis, but also to screen for inhibitors of protein-protein interactions. This is demonstrated in Chapter 3 by the discovery of inhibitors of the interactions between the human oncoprotein MDM2 and p53.

The combined application of computer simulations and NMR spectroscopy to find inhibitors of the IGF-I and IGF-binding protein 5 interaction is discussed in Chapter 4.

For heteronuclear NMR spectroscopy labeled protein samples are needed. A novel medium for expression of selectively ^{15}N labeled proteins in SF9 insect cells was recently invented in our group. The role NMR spectroscopy has played in understanding the metabolism of these cells is shown in Chapter 5.

The royal discipline of biochemical NMR spectroscopy remains the solution of the tertiary structure of medium sized proteins. In Chapter 6 work on the determinaton of the solution structure of the adenylyl cyclase associated protein (CAP) from *Dictyostelium discoideum* is reported.

Finally, Chapter 7 describes the NMR characterization of the green fluorescent protein, which paved the way for investigations of its dynamical properties.

Chapter 2

Application of NMR in structural proteomics

2.1 Introduction

In the time of structural proteomics when protein structures are targeted on a genome-wide scale the detection of "well-behaved" proteins that would yield good quality NMR spectra or X-ray images is the key to high-throughput structure determination. Already simple one-dimensional proton NMR spectra provide enough information for assessing the folding properties of proteins. Heteronuclear two-dimensional spectra are routinely used for screenings that reveal structural as well as binding properties of proteins. NMR thus can provide important information for optimizing conditions for protein constructs that are amenable to structural studies.
In this chapter an overview of the applications of NMR in screening for protein samples that are suitable for structure elucidation by both NMR spectroscopy and X-ray crystallography is given. These applications could be the main contribution of NMR to structural proteomics as securing "well-behaved" samples is expected to be the rate-determining step in any structural proteomics project (Christendat et al., 2000).

2.2 Screening for proteins amenable to structural analysis

It has been widely assumed that nuclear magnetic resonance spectroscopy will play an important role in structural proteomics complementing X-ray crystallography for small and medium size proteins (below 30 kDa) (Montelione et al., 2000; Prestegard et al., 2001). About 17%

of the structures deposited in the Protein Data Bank (PDB) have been solved by NMR spectroscopy, most of which do not have corresponding crystal structures (Prestegard et al., 2001; Sali, 1998). Even if NMR will remain a "poor daughter" of the X-ray method in determining structures of large proteins, it nevertheless can deliver strong results in several areas of structural biochemistry. It is the basis for a wide range of experiments to determine structure-function relationships, to find binding partners with their specific binding sites (Shuker et al., 1996), to investigate dynamics of proteins (Renner & Holak, 2001), to distinguish multiple conformations (Mühlhahn et al., 1998), to compare apo and holo forms of proteins and map the binding sites of their cofactors (Wijesinha-Bettoni et al., 2001) or to determine pKa values of ionizable groups (Fielding, 2000), to name just a few. A series of spectra taken under different conditions may be used to monitor aggregation and even formation of amyloid fibrils (Zurdo et al., 2001), to determine K_D values of binding partners (Shuker et al., 1996) or to track hydrogen exchange with real time NMR in proteins dissolved in D_2O (Canet et al., 2002). The ability to detect ligands binding only very weakly to target molecules has made NMR also increasingly important in drug discovery (Diercks et al., 2001; Pellecchia et al., 2002).

In a recent investigation of roughly 500 proteins from the genome of a single organism Christendat et al. (2000) found that only 10-15% of these proteins yielded samples that where of sufficient quality for structural analyses by either NMR spectroscopy or X-ray crystallography. Clearly a method to screen for these "well-behaved" proteins as well as to optimize the samples of the vast rest is needed.

The unique strength of NMR lies in its capability to semi-quantitatively estimate unstructured regions of the polypeptide chain in the otherwise partially folded protein and to identify proteins that are heterogeneous because of aggregation or other conformational effects. The various applications of NMR in structural proteomics will be illustrated by typical examples in the following sections.

2.3 One-dimensional NMR

A simple one-dimensional proton experiment, the most basic spectrum in NMR spectroscopy that can be acquired in a short time (usually not longer than a few minutes) for samples as dilute as 0.01 mM contains already a great amount of information. The lower panel of Figure 2.1 shows an example of an unfolded protein with a large and broad signal at approximately 8.3 ppm. An unfolded protein shows a small dispersion of the amide backbone chemical shifts

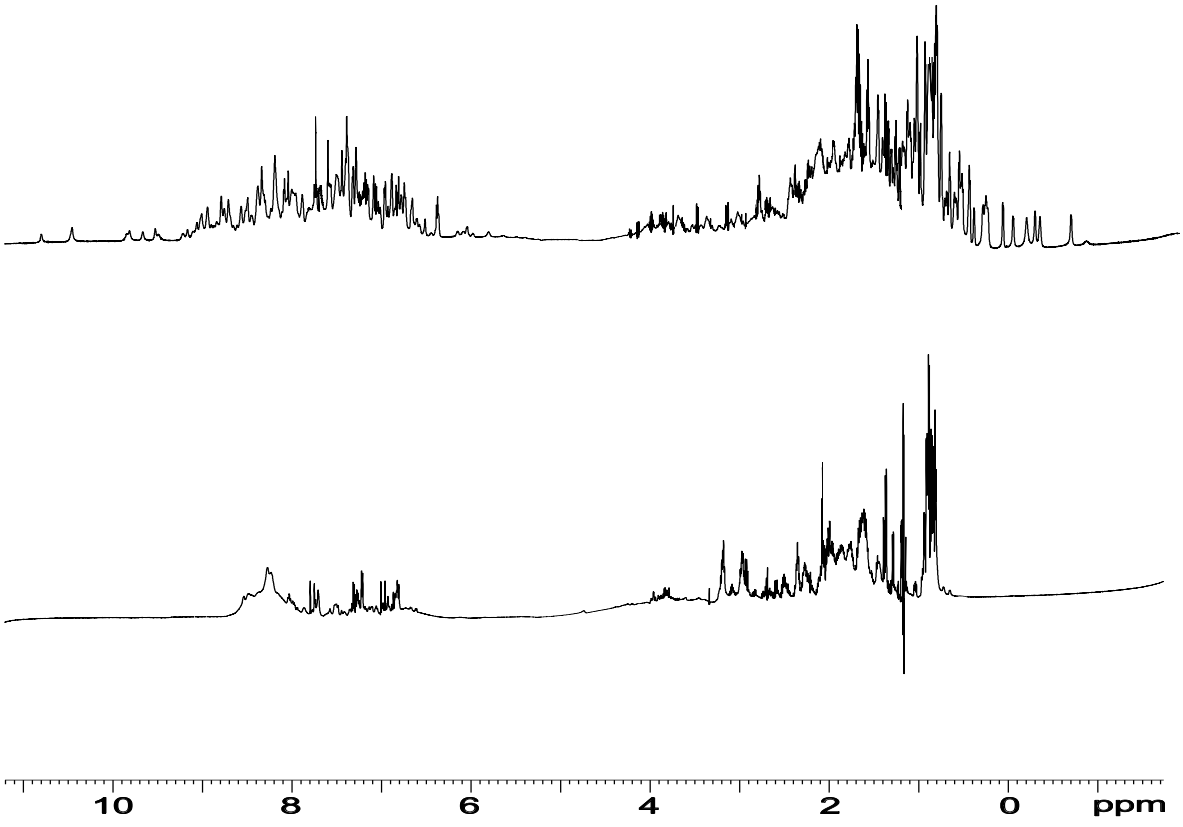

Figure 2.1: Characterization of protein structures by one-dimensional NMR spectroscopy: Upper panel: A typical one-dimensional proton NMR spectrum of a folded protein with signal dispersion downfield (left) of 8.5 ppm and upfield (to the right) of 1 ppm. Spectra show the N-terminal 176 residue domain of the cyclase associated protein (CAP) at pH 7.3. Lower panel: An unfolded protein sample. Strong signals appear around 8.3 ppm, the region characteristic for amide groups in random coil conformation. No signal dispersion is visible below approximately 8.5 ppm. Also to the right of the strong methyl peak at 0.8 ppm no further signals show up. The sample is an unfolded domain of the IGF binding protein 4 (IGFBP-4, residues 147 - 229).

(Wüthrich, 1986). Particularly the appearance of intensities at chemical shifts near 8.3 ppm is an excellent indicator for a disordered protein, as this is a region characteristic of backbone amides in random coil configuration. On the other hand signal dispersion beyond 8.5 ppm (8.5 - 11 ppm) proves a protein to be folded. Due to the different chemical environment and thus the varying shielding effects the resonances of the single protons will be distributed over a wide range of frequencies. A typical intensity pattern of a folded protein is shown in the upper panel of Figure 2.1. Following the same argument, in the aliphatic region of the spectrum between 1.0 and -1.0 ppm a large signal dispersion versus a steep flank of the dominant peaks at approximately 1 ppm separates a structured from an unfolded protein (Figure 2.1 upper and lower panel, respectively).

Close inspection of one-dimensional spectra will also yield quantitative information on the extent of folding in partially structured proteins or their domains. Figure 2.2 shows two spectra of a 20 kDa protein. In the upper spectrum a mixture of approximately 50% folded and 50% unfolded protein can be identified by observing both the signal dispersion and the prominent peak at 8.3 ppm. The lower spectrum shows the same sample after removal of the unfolded macromolecules by gel filtration. The "random-coil peak" disappeared and the signal pattern is that of a completely structured protein.

While the signal dispersion of the resonances is generally connected to folding, aggregation can be detected by observing the line width of the signals. Due to faster relaxation mechanisms, the NMR signal from larger molecules will decay much faster than that from smaller ones (Abragam, 1961). This in turn will produce broader lines for the resonances of larger molecules. Thus the line widths of the signals in any NMR spectrum are correlated to the size of the molecule. Both these aspects may be appreciated in Figure 2.3. The upper panel shows the spectrum of the 246 residue IGFBP-5 (Kalus et al., 1998) that exhibits a rather large peak at the random-coil value of 8.3 ppm and some signals down-field (that is shifted to higher ppm values) close to the noise level. The IGFBP-5 protein comprises conserved N- and C-terminal domains of 90 and 112 amino acids respectively and a central domain of 40 amino acids. Spectra of the C-and N-terminal fragments of the same protein (Figure 2.3 middle and lower panel respectively) show, that the unstructured regions are located in the C-terminal fragment. The N-terminal fragment shows nice signal dispersion, characteristic of a structured protein. Again in this example the quantitative information on the extent of folding that is available from 1D-NMR can be appreciated. The full-length protein is only to about 50-60% folded. This spectrum (Figure 2.3 upper panel) may be seen as a superposition of the two other spectra which show

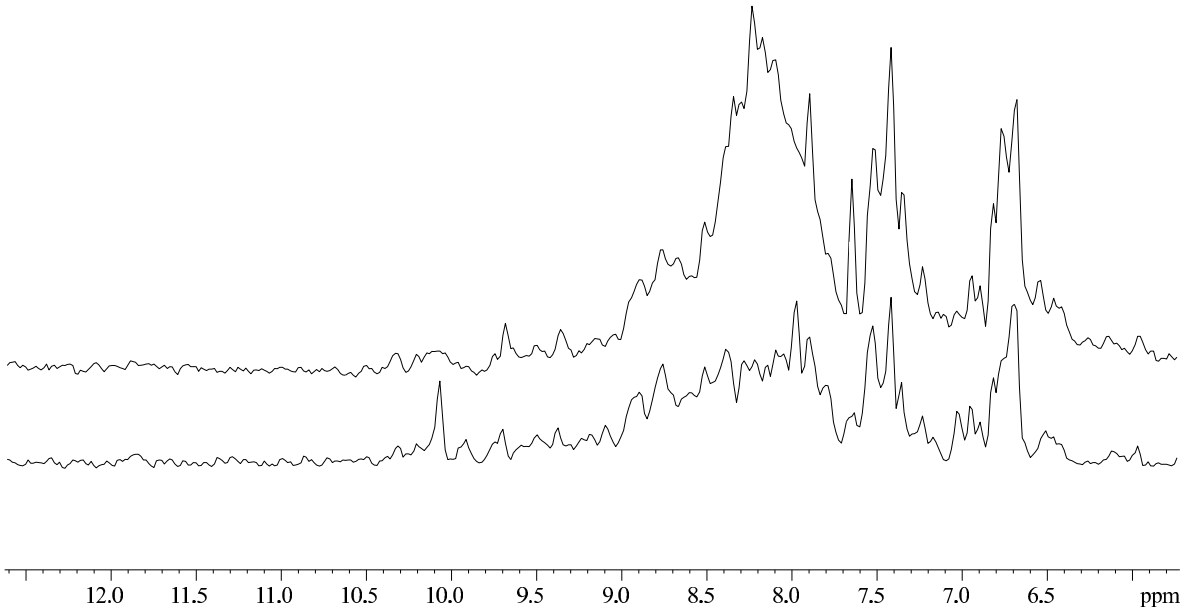

Figure 2.2: The amide region of a 20 kDa protein: The upper trace shows a 1:1 mixture of folded and unfolded proteins, the lower trace the same sample after removing the unfolded proteins by gel filtration.

the C-terminal fragment which is to about 30% folded and the fully folded N-terminal fragment (middle and lower panel respectively, the rest of the unstructured residues originate from the central domain of IGFBP-5).

Note that also the line width of the individual signals has improved dramatically in the smaller fragments. Using the line width from known monomeric proteins of a given size as a reference, observation of the line width in a one-dimensional spectrum will also yield information on the molecular weight and aggregation of the molecule under investigation. Furthermore, attempts to prevent aggregation by e.g. dilution of the sample, addition of mild detergents as CHAPS or lowering the pH value can thus be monitored by NMR to find optimal sample conditions (Kalus et al., 1998; Anglister et al., 1993; Edwards et al., 2000).

While the extent of folding is crucial both for X-ray crystallography and NMR, aggregation is not. Actually some proteins that yield rather poor NMR spectra due to aggregation or low solubility might give excellent crystals as was the case for p19^{INK4d} (Kalus et al., 1997; Baumgartner et al., 1998). Thus sample conditions that are optimal for crystallography might not necessarily be optimal for NMR spectroscopy and vice versa. This fact does not however reduce the value of insights given by NMR for crystallography.

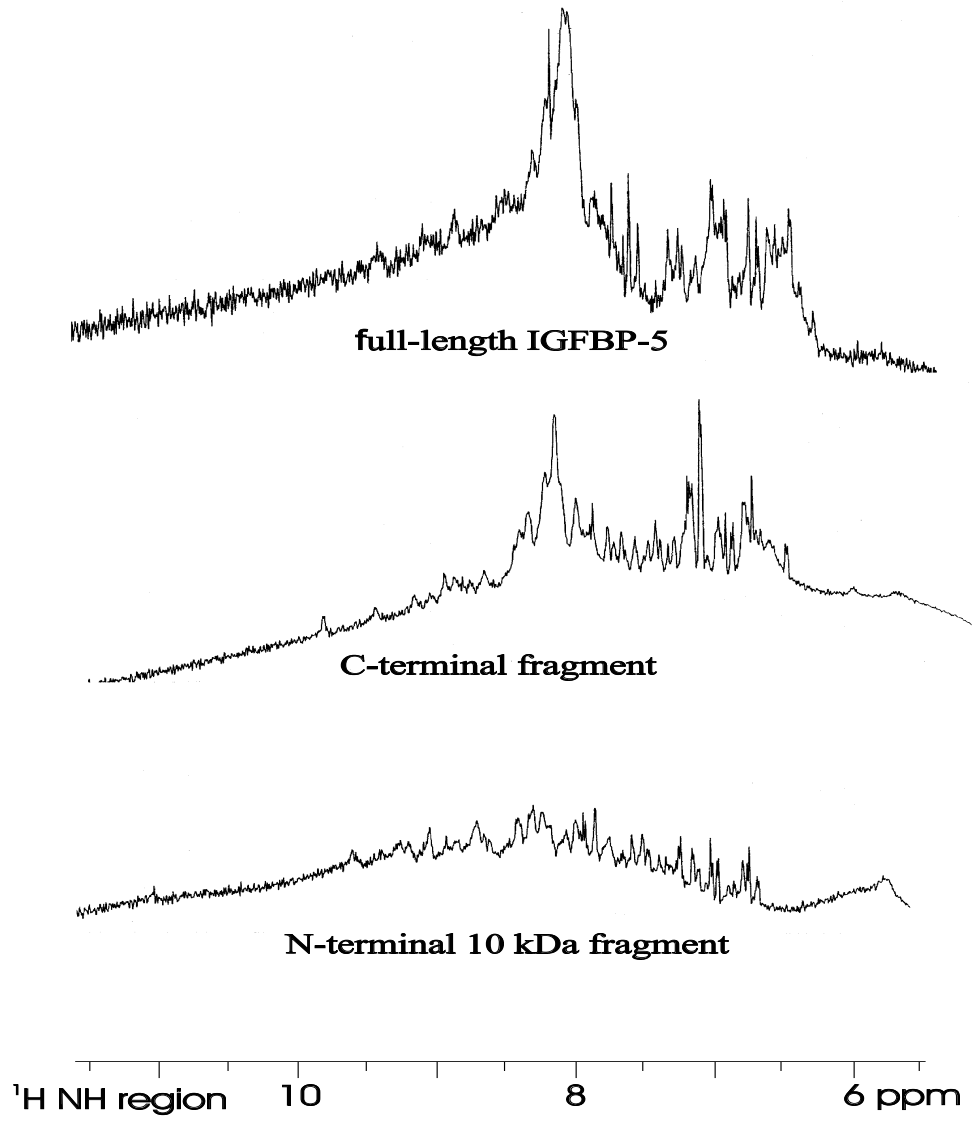

Figure 2.3: Amide region of one-dimensional spectra of the IGF binding protein-5: The upper, middle and lower panels show the full-length protein of 246 amino acid residues, a C-terminal fragment of 112 residues and a N-terminal fragment of 94 residues, respectively.

Comparing Figures 2.2 and 2.3 it has to be pointed out that distinguishing between a protein that is only partially folded and a mixture of folded and unfolded proteins is difficult with NMR without having additional information from e.g. gel filtration or other biochemical methods.

One-dimensional spectra may additionally provide information on α-helical or β-strand structures in a protein. The C^{α} protons in a helix display few resonances in the region between 5 and 6 ppm, while those in a β-sheet resonate in this region (Wishart et al., 1991).

The use of one-dimensional spectra to screen for optimal, fully folded protein fragments may be illustrated by the example of the cyclase associated protein (CAP) (Gottwald et al., 1996, see also chapter 6). The initial construct of the protein comprising 226 amino acids showed considerable line width and would not crystallize (Figure 2.4 A). After the sample was left at room temperature for 7 days, another one-dimensional spectrum showed, on the one hand, degraded peptide fragments but on the other hand still broad lines not in agreement with the expected shorter protein fragment (Figure 2.4 B). Mass spectrometry revealed the presence of several protein fragments of different length ranging from 226 to 173 amino acids. Thus the very sharp peaks around 1 ppm could be attributed to the cleaved peptide fragments, while the line width corresponds to the overlap of slightly varying resonances from several fragments of different length. Based on these results a new fragment of the protein of 176 residues which contained only the unchanged core of the protein was cloned and expressed in E. coli. This fragment was not further degraded even after several months. The spectrum of this protein fragment is shown in Figure 2.4 C. Note the absence of the sharp resonances and the superior line width. NMR has revealed a stable folded core of the protein that was then successfully subjected to the NMR and X-ray structure analysis.

The prominent signals from the small peptide fragments provide also an example for the examination of a sample's purity. Any small compounds, be it peptides or other impurities will readily show in a one-dimensional spectrum.

2.4 Two-dimensional NMR

Due to the greatly improved resolution of two-dimensional experiments, these are frequently used for screening and binding studies. The simplest and most powerful among them is the heteronuclear single-quantum coherence (HSQC) experiment. In a large scale approach Yee et al. (2002) have recently investigated more than 500 proteins from five different organisms, using ^{15}N-HSQC experiments to screen for those proteins amenable for NMR structure analysis.

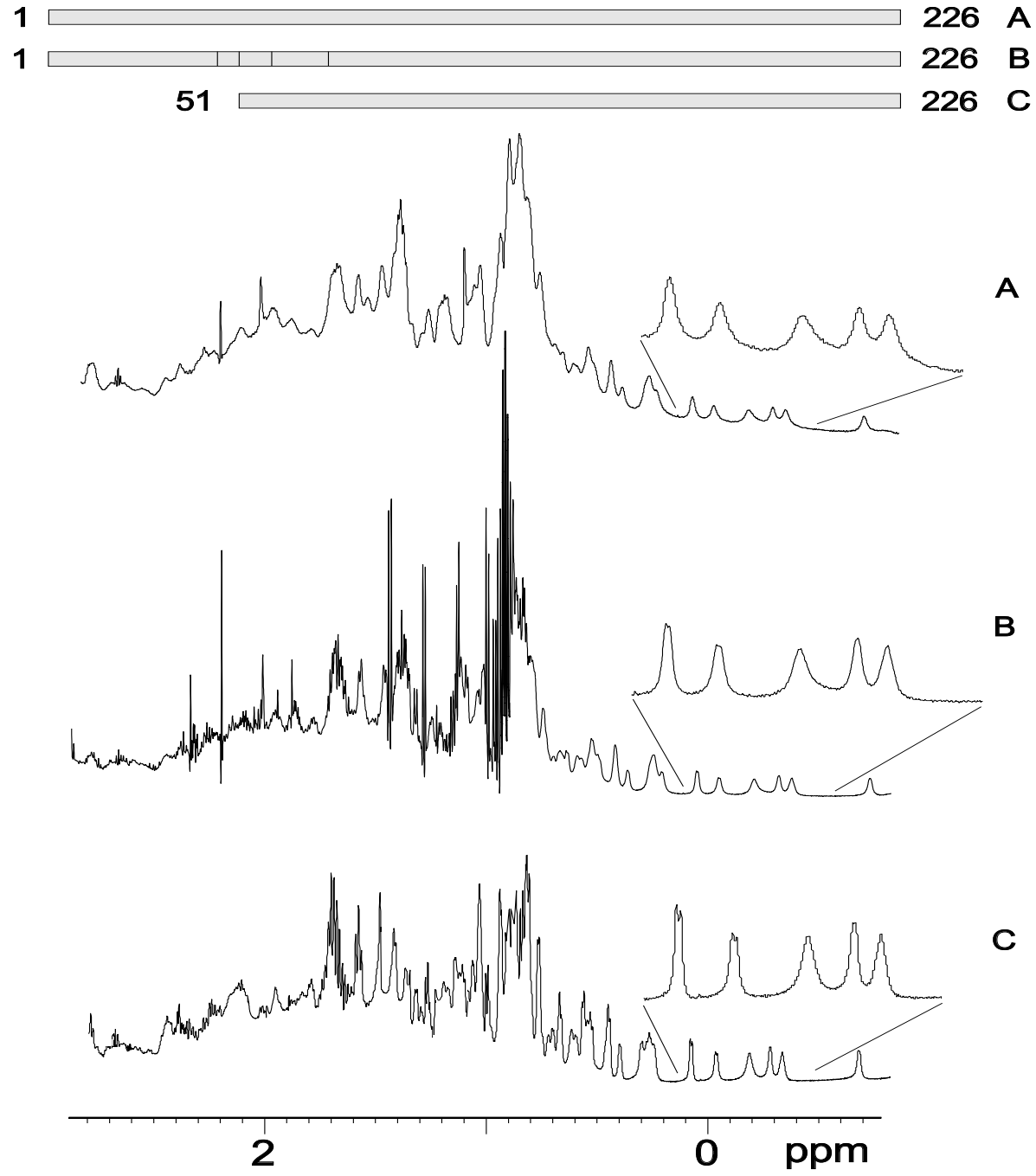

Figure 2.4: The aliphatic region of one-dimensional spectra of the cyclase associated protein (CAP): A) The N-terminal 226 residue construct as indicated in the scheme above. B) A mixture of several constructs of different length (residues 1-226, 44-226, 51-226, 56-226). The overlap leads to broader lines compared to spectrum A). Short peptides give rise to sharp signals around 1 ppm. C) The stable core of the protein (residues 51-226) only. The sharp signals from impurities are removed and the linewidth is substantialy improved compared to spectrum A).

This spectrum is the first step in any structure elucidation as it maps the backbone amide groups of a protein according to their proton and nitrogen frequencies. A whole set of three-dimensional spectra later used to assign the NMR signals to their respective amino acid residue is based on the HSQC experiment. For this kind of spectrum ^{15}N-labeled protein samples are required. The HSQC shows one peak for every proton bound directly to a nitrogen atom and thus exactly one signal per residue in the protein (apart from proline which is devoid of proton bound nitrogen and some additional side chain signals appear which can easily be identified).

The positions of the peaks are indicative of structured or disordered proteins in the same way as described above for the one-dimensional spectrum (Figure 2.5). In the spectrum of an unfolded protein all signals cluster in a characteristic "blob" around a 1H frequency of 8.3 ppm with little signal dispersion in both dimensions. In the spectrum of a structured protein, the peaks show large signal dispersion. Thus if the peaks are assigned their respective sequential position in the polypeptide chain, disordered regions may be identified.

As the number of signals in the HSQC spectrum corresponds approximately to the number of residues in the protein under investigation, conformational heterogeneity can easily be detected by a surplus of peaks. To optimize sample conditions pH titrations or titrations with cofactors or other molecules as well as variation of temperature may be performed while repeatedly recording HSQC spectra. This is feasible since the NMR method is non-destructive and experiments may be repeated several times. It has, for example, been shown by NMR observed titrations that low temperatures and neutral pH stabilize the folded state of a SH3 domain of Drosophila drk, while high temperatures and low pH tend to favor the unfolded state (Zhang & Forman-Kay, 1995). On the other hand low pH has also been reported to prevent aggregation as observed by line width comparison (Anglister et al., 1993).

For full NMR structure investigations samples of 200 - 400 μl with a protein concentration of 0.5-1.0 mM to are required. This corresponds to about 10 - 15 mg/ml of the protein which is the concentration usually used for crystallographic screening. Spectra can also be recorded in up to 1 mM Tris buffer. Note again that NMR does not destroy the sample, so it is possible to continue with crystallization attempts after NMR characterization.

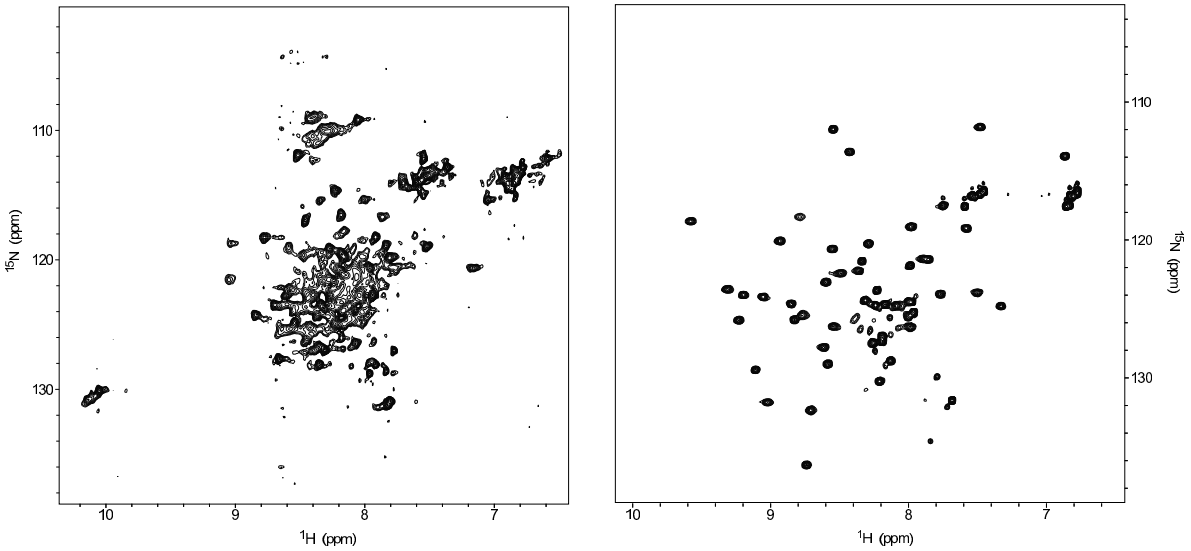

Figure 2.5: ^{15}N-HSQC spectra of unfolded and folded proteins. The left panel shows a ^{15}N-HSQC spectrum of partially unstructured protein fragment of 80 amino acid residues. All signals cluster around a ^{1}H frequency of 8.3 ppm. Also the signal dispersion in the ^{15}N dimension is limited. The broad unresolved signals in the middle of the spectrum indicate either aggregation in the sample or conformational heterogeneity on a ms-μs timescale (both cases being unfavorable for NMR studies). The signal at 10 ppm is not diagnostic for a folded protein, but stems from the sidechain amide group of a tryptophan residue. The right panel shows the spectrum of a folded, 55 residue long construct of the IGFBP-5 protein. The peaks show a large signal dispersion in both dimensions.

Chapter 3

Chalcone Derivatives Antagonize Interactions between the Human Oncoprotein MDM2 and p53

3.1 Introduction

The oncoprotein MDM2 (human murine double minute clone 2 protein) inhibits the tumor suppressor protein p53 by binding to the p53 transactivation domain. The p53 gene is inactivated in many human tumors either by mutations or by binding to oncogenic proteins. In some tumors, such as soft tissue sarcomas, overexpression of MDM2 inactivates an otherwise intact p53, disabling the genome integrity checkpoint and allowing cell cycle progression of defective cells. Disruption of the MDM2/p53 interaction leads to increased p53 levels and restored p53 transcriptional activity, indicating restoration of the genome integrity check and therapeutic potential for MDM2/p53 binding antagonists. In this chapter it is shown by multidimensional NMR spectroscopy that chalcones (1,3-diphenyl-2-propen-1-ones) are MDM2 inhibitors that bind to a subsite of the p53 binding cleft of human MDM2. Biochemical experiments showed that these compounds can disrupt the MDM2/p53 protein complex, releasing p53 from both the p53/MDM2 and DNA-bound p53/MDM2 complexes. These results thus offer a starting basis for structure-based drug design of cancer therapeutics.

3.2 Biological Context

Amplification of the MDM2 gene is observed in a variety of human tumors (Juven-Gershon & Oren, 1999; Momand et al., 1998). MDM2 is an oncogene product that binds to the transactivation domain of the p53 tumor suppressor protein (Lane & Hall, 1997; Oliner et al., 1992; Lozano & Montes de Oca Luna, 1998; Kussie et al., 1996). By binding to p53, MDM2 inhibits the ability of p53 to activate transcription (Oliner et al., 1993) and promotes the rapid degradation of p53 (Haupt et al., 1997; Kubbutat et al., 1997). Increasing MDM2 levels thus raises the signal threshold necessary for p53-induced apoptosis (Oliner et al., 1993; Haupt et al., 1997; Kubbutat et al., 1997; Momand et al., 1992; Midgley & Lane, 1997) and retards the rate of the p53-induced expression of the cell cycle inhibitor p21 (Momand et al., 1992; Chen et al., 1993). Studies comparing MDM2 overexpression and p53 mutation concluded that these are mutually exclusive events, supporting the notion that the primary impact of MDM2 amplification in cancer cells is the inactivation of the resident wild-type p53 (Juven-Gershon & Oren, 1999; Momand et al., 1998; Oliner et al., 1993). It has been shown recently that a peptide homologue of p53 is sufficient to induce p53-dependent cell death in cells overexpressing MDM2 (Wasylyk et al., 1999). This result provides clear evidence that disruption of the p53/MDM2 complex might be effective in cancer therapy. Chalcone derivatives (compounds derived from 1,3- diphenyl-2-propen-1-one) have been described in the literature as inhibitors of chemoresistance (Daskiewicz et al., 1999), ovarian cancer cell proliferation (Devincenzo et al., 1995), pulmonary carcinogenesis (Wattenberg, 1995), proliferation of HGC-27 cells derived from human gastric cancer, and other tumorigenic effects (Shibata, 1994). Licochalcone-A, a chalcone derivative found in the licorice root, has been associated with a wide variety of anticancer effects, along with other potential benefits (Park et al., 1998).

3.3 Ligand Binding

Determination of binding sites of lead chalcone compounds (Figure 3.1) were carried out using ^{15}N-HSQC NMR spectroscopy of the ^{15}N isotopically enriched domain of human MDM2 including residues 1-118. A nearly complete assignment of the backbone ^{1}H and ^{15}N NMR resonances was obtained for the uncomplexed MDM2 previously (Stoll et al., 2000, see also Figure 3.2). The NMR ^{15}N-{^{1}H} NOE experiment indicated that the folded core of the MDM2 domain in solution extends from T26 to N111 (Figure 3.3). This is in good agreement with the

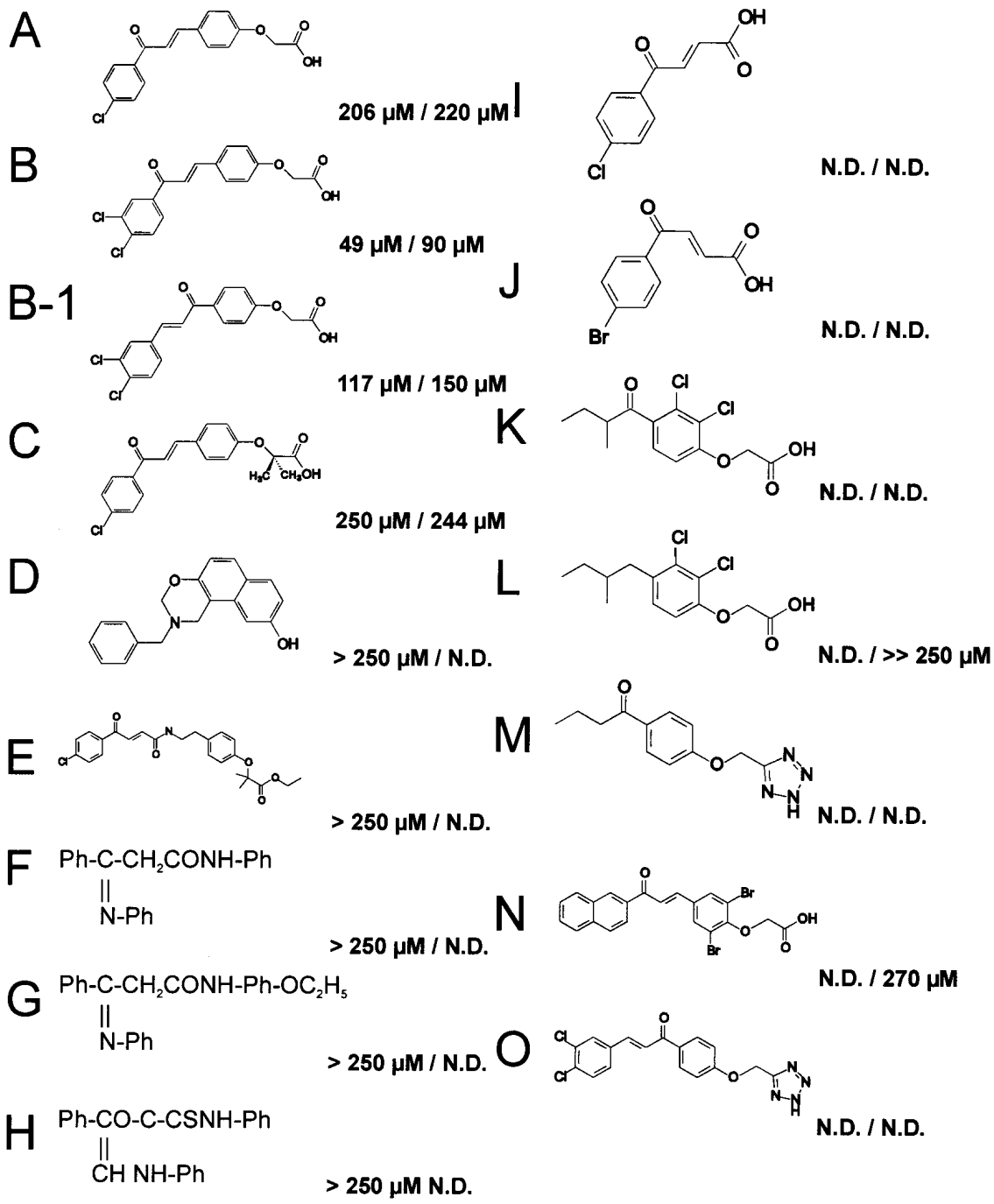

Figure 3.1: A representative collection of basic chalcone skeletons used in this study. Inhibition of MDM2 binding to p53 measured by ELISA (IC_{50} values given on the left side of the slash) and by NMR titration experiments (K_D values given on the right side of the slash). Compound D was studied as a negative control.

crystal structures of N-terminal domains of human and Xenopus MDM2 in complex with a transactivation domain peptide of p53, where the MDM2 structure was also defined from T26 to V109 (Kussie et al., 1996). The p53 peptide, comprising the residues 15 to 29, binds to an elongated hydrophobic cleft of the MDM2 domain. The interaction is primarily hydrophobic in character; only two hydrogen bonds are found between MDM2 and the p53 peptide. The hydrophobic surfaces of MDM2 and p53 are sterically complementary at the interface. The binding surface of p53 is dominated by a triad of p53 amino acids (F19, W23, and L26) that bind along the MDM2 cleft and define the corresponding phenylalanine, tryptophan, and leucine subpockets for the p53/MDM2 interaction (Kussie et al., 1996). In this classification, the leucine pocket is defined by Y100, T101, and V53, the tryptophan pocket is defined by S92, V93, L54, G58, Y60, V93, and F91, the phenylalanine pocket is defined by R65, Y67, E69, H73, I74, V75, M62, and V93 (Kussie et al., 1996). As a control experiment using a known stable MDM2/ inhibitor complex, MDM2 was titrated with the p53 peptide comprising residues E17 to N29 (Figure 3.3, panel A). NMR spectra showed that the p53 peptide/MDM2 complex was long-lived on the NMR chemical shift time scale. This is in agreement with ELISA data that showed an apparent K_D of 0.6 μM (Kussie et al., 1996). As can be seen in Figure 3.3, panel A, almost all amino acids of the free MDM2 exhibit changes in chemical shifts upon complexation with the p53 peptide. The analysis of ligand-induced $^{1}H^{N}$ and the ^{15}N shifts was performed by applying the equation of Pythagoras to weighted chemical shifts which is in concordance with the recent literature (Pellechia et al., 1999). The largest shifts lined the three binding subpockets of p53 on MDM2 (Figure 3.3, panel A). The full set of MDM2/p53 interface residues comprises M50, L54, L57, G58, I61, M62, Y67, H73, V75, F91, V93, H96, I99, and Y100 of MDM2 (Kussie et al., 1996). Additionally, significant shifts are observed for β-strand residues T26, L27, V28, R29, L107, and V108 and for residues L34, L37, and K64. Shifts observed for amides outside the binding regions may be caused by secondary effects, such as allostery or change in mobility upon binding, and do not necessarily indicate direct binding of the p53 peptide to MDM2. Such possible secondary effects (e.g., residues L34, L37, and K64) must be considered when analyzing ligand binding to allosteric proteins.

All K_D values determined by NMR spectroscopy fully agree with the affinities measured by the ELISA binding assay (see Figure 3.1). Compound A, with an ELISA IC_{50} value of 206 μM, shows the strongest shifts at the peptide groups of E52, V53, L54, F55, Y56, L57, G58, Y60, I61, and H73 (Figure 3.3). Except for H73, all of these are found on the α-helix comprising residues M50-R65; the H73 shift is attributed to secondary or allosteric effects. The shift pat-

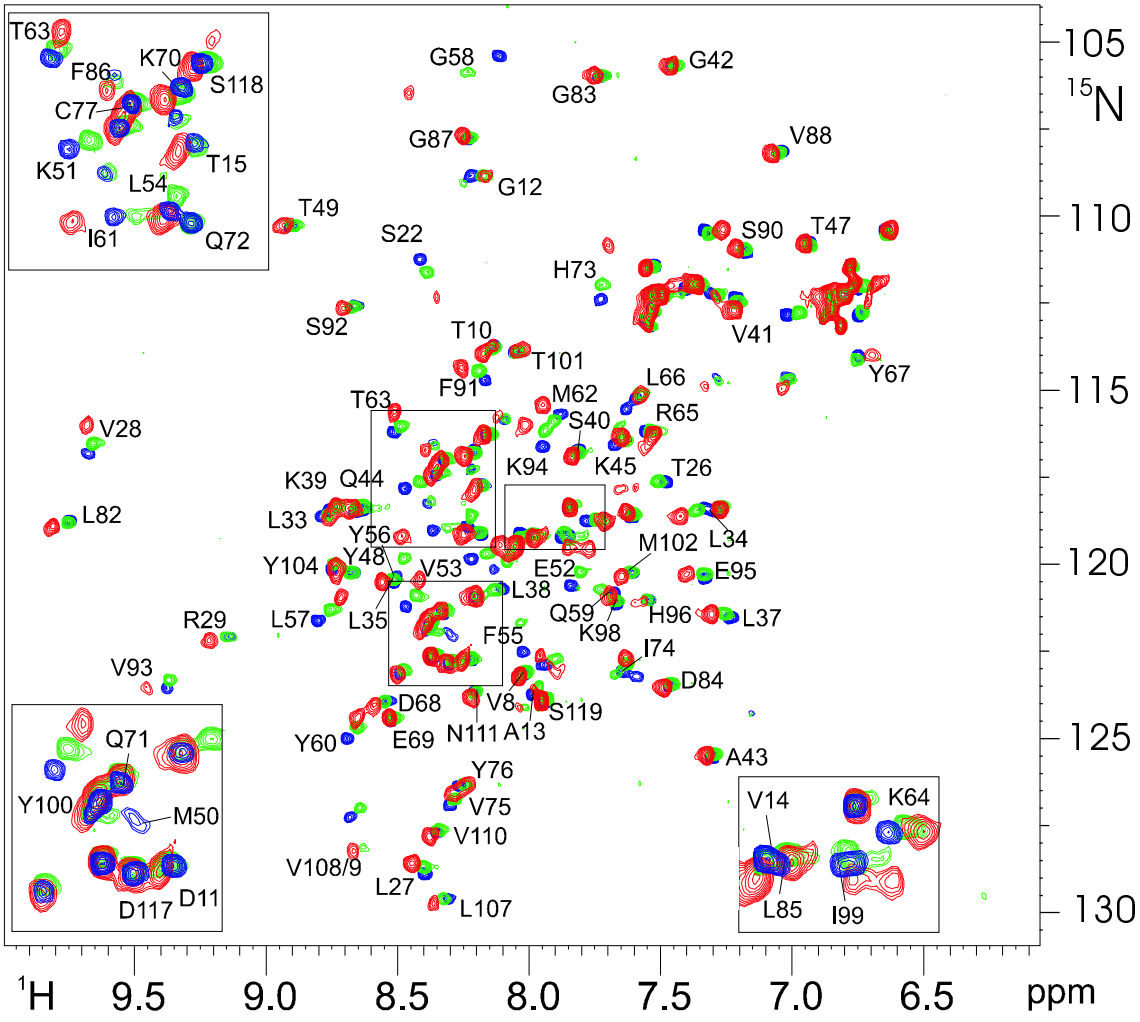

Figure 3.2: 500 MHz 2D ^{1}H-^{15}N HSQC spectrum of human MDM2 titrated with increasing amounts of chalcone C. Cross-peaks for apo- MDM2 are marked in blue; green and red cross-peaks indicate 50 and 100% complexation of MDM2 by chalcone C. Residue specific assignment of the backbone ^{1}H and ^{15}N frequencies is indicated.

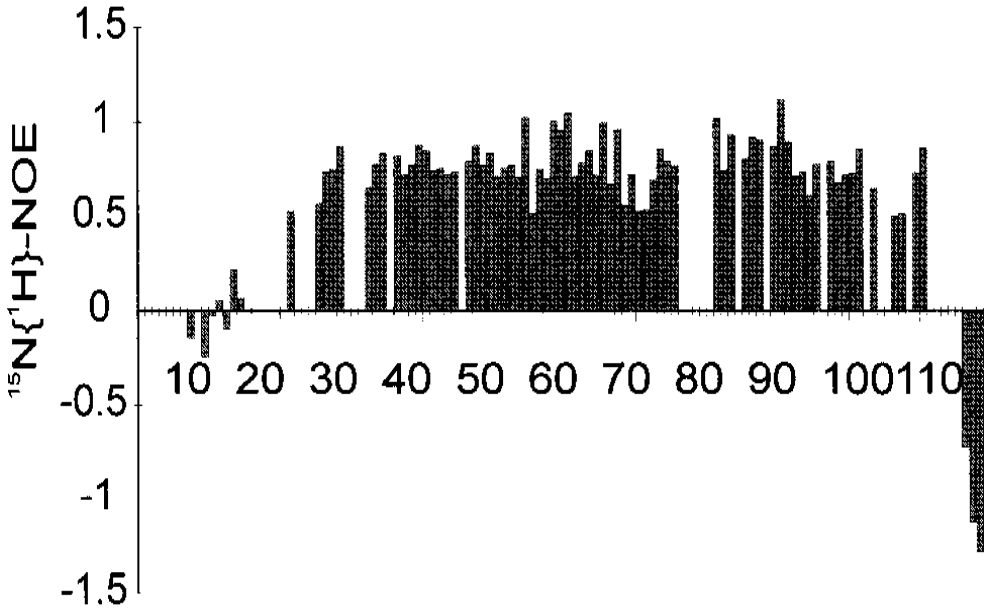

Figure 3.3: $^{15}N\{^{1}H\}$-NOE for the backbone amides of human MDM2. Residues for which no results are shown correspond either to prolines or to residues where relaxation data could not be extracted.

tern is consistent with binding in the tryptophan pocket of MDM2. Compounds B and B-1 yielded similar chemical shift patterns as compared to compound A (Figure 3.3). The shifts observed for compounds B and B-1 cannot reliably be used to localize the inhibitor interaction site because these inhibitors induce precipitating MDM2/MDM2 interactions that also contribute to the chemical shift pattern. The same is true for compounds N and O. Chalcone C differs from A by the addition of two methyl groups near the acid terminus, an alteration that insignificantly affects the IC_{50} value (250 μM). The overall NMR shift perturbation pattern is similar to that observed for chalcone A (Figures 3.1 and 3.3). The detailed shift perturbation pattern, however, is changed by the dimethyl substitution: the perturbations observed for T26, K51, and E52 are new or greater, while the perturbations at Y56 and I61 caused by compound C are weakened (Figure 3.3, panels B and E).

In conclusion, it could been shown that chalcone derivatives bind to the tryptophan pocket of the p53 binding site of MDM2 and are able to dissociate the p53/MDM2 complexes. Therefore chalcones, as antagonists of the p53/MDM2 interaction, offer the starting point for structure-based drug design for cancer therapeutics in strategies that abolish constitutive inhibition of p53 in tumors with elevated levels of MDM2 or, more generally, in strategies that enhance p53

activity.

3.4 NMR Spectroscopy

NMR measurements consisted of monitoring changes in chemical shifts and line widths of the backbone amide resonances of uniformly ^{15}N-enriched MDM2 samples (Shuker et al., 1996; McAlister et al., 1996) in a series of HSQC spectra as a function of a ligand concentration. No changes in chemical shifts were observed between samples of different concentrations (0.03-0.5 mM) and pH values between 6.5 and 7.5. For titration experiments, 0.1-0.3 mM of human MDM2 in 50 mM KH_2PO_4, 50 mM Na_2HPO_4, 150 mM NaCl, pH 7.4, and 5 mM DTT was used. The chalcone derivatives were lyophilized and finally dissolved in DMSO-d6. No shifts were observed in the presence of 1% DMSO (the maximum concentration of DMSO in all NMR experiments after addition of inhibitors). All chalcone-MDM2 complexes showed a continuous movement of several NMR peaks upon addition of increasing amounts of inhibitors. From these experiments, the spectra of MDM2 could be assigned unambiguously. The complexes of human MDM2 and the chalcones were prepared by mixing the protein and the ligand in the NMR tube. Typically, NMR spectra were recorded 15 min after mixing at room temperature. An initial screening of all compounds used in this study was performed with a 10-fold molar excess of chalcone to human MDM2. All subsequent titrations were carried out until no further shifts were observed in the spectra. Saturating conditions were achieved at a molar ratio of chalcone to MDM2 of 6 for chalcone A, of 2 for chalcone B, of 2 for chalcone B-1, and of 6 for chalcone C, for example. Typically, the concentration of human MDM2 was 0.1 mM and the final concentration of the chalcone ligand was 50 mM in each titration. All K_D values obtained by NMR spectroscopy are based on at least six data points. From the independently determined IC_{50} values and the K_D constants, one ligand binding site for these chalcones per MDM2 is calculated taking into account the molar ratio of ligand to protein in the NMR experiments. Quantitative analysis of induced chemical shifts were performed on the basis of spectra obtained at saturating conditions of each chalcone. Analysis of ligand-induced shifts was performed by applying the equation of Pythagoras to weighted chemical shifts: $\Delta\delta_c(^1\mathrm{H},^{15}\mathrm{N}) = [\{\mid \Delta\delta(^1\mathrm{H}) \mid^2 + 0.2\cdot \mid \Delta\delta(^{15}\mathrm{N}) \mid^2\}^{0.5}]$. The p53 peptide/MDM2 complex was long-lived on the NMR chemical shift time scale (lifetimes $\gg$ 0.2 ms) (Wüthrich, 1986). Two separate sets of resonances were observed in the ^{1}H-^{15}N HSQC spectra, one corresponding to free MDM2 and the other to MDM2 bound to the p53 peptide. For well-resolved, isolated peaks, the assignment of Figure 3.2 could be transferred to

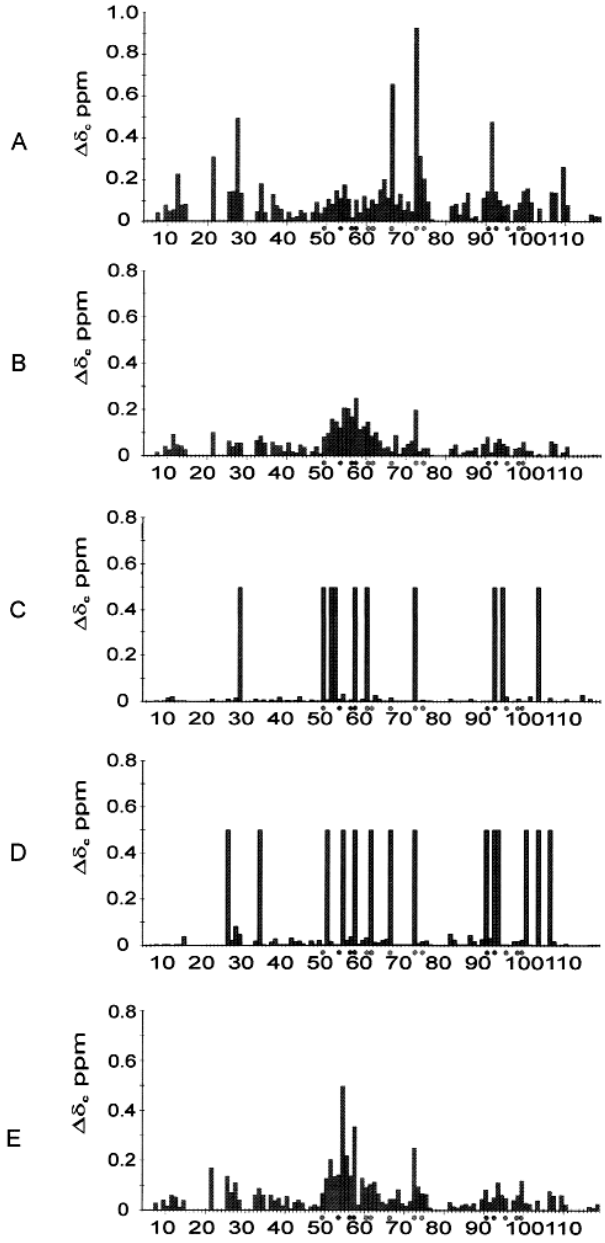

Figure 3.4: Plots of induced differences in the NMR chemical shifts versus the amino acid sequence. (A) The p53 peptide; (B) inhibitor A; (C) inhibitor B; (D) inhibitor B-1 (for the maximum induced shifts for B and B-1 see explanation in experimental procedures); (E) inhibitor C. Dots mark the leucine-, tryptophan-, and the phenylalanine-binding site on human MDM2.

the resonances in the peptide complex (54% of all backbone amide resonances in the ^{1}H-^{15}N HSQC). For the rest of the shifts, assignment of $\Delta\delta_c(^{1}\text{H},^{15}\text{N})$ upon complex formation was carried out in a conservative manner, i.e., for these shifts the distance in ppm to the closest peak in complexed MDM2 was chosen. In addition, all selectively enriched samples of human MDM2 (^{15}N-Val, ^{15}N-Leu, ^{15}N-Phe, and reverse ^{14}N-His) were titrated with the p53 peptide to confirm a subset of MDM2/p53 complex assignments. Only $\Delta\delta_c(^{1}\text{H},^{15}\text{N})$ values larger than 0.1 ppm were considered to be significant. $\Delta\delta_c(^{1}\text{H},^{15}\text{N})$ smaller than 0.1 ppm were found for 37 residues. Erroneous conclusions could result if some of the residues with $\Delta\delta_c(^{1}\text{H},^{15}\text{N}) < 0.1$ ppm were actually in contact with the inhibitor. However, the internal consistency of our results corroborates our analysis; for example, no core buried residue was found that had $\Delta\delta_c(^{1}\text{H},^{15}\text{N}) > 0.1$ ppm. Furthermore, all residues of human MDM2 involved in binding to the p53 peptide also show significant shifts $\Delta\delta_c(^{1}\text{H},^{15}\text{N})$ upon complexation with the peptide (Kussie et al., 1996). For compounds B and B-1 (Figure 3.3, panels C and D), the maximum shifts shown at $\Delta\delta_c =$ 0.5 ppm correspond to the cross-peaks of the folded core of MDM2 whose line-widths broaden 2-fold upon addition of either B or B-1 in the molar ratio of B-1 to MDM2 1:1 and disappear thereafter at the titration ratio 2:1 (McAlister et al., 1996). Compound D (Figure 3.1) was studied as a negative control because it did not inhibit MDM2 binding to a p53 peptide as measured by ELISA. This compound does not bind to apo-MDM2, as no ^{1}H and ^{15}N shifts greater than 0.1 ppm were observed in the NMR spectra. As this compound was available in our laboratory and because of its similar size as compared to the chalcone skeleton, we have selected this heterocyclic system as a negative control for any organic compound. Other negative control NMR titration experiments included the chemically synthesized chromophore of the green fluorescent protein as well as a synthetic 22-residue peptide. None of the control ligands led to significant chemical shift perturbations (data not shown). Chalcone B-1 generally enhances the intrinsic tendency of MDM2 to aggregate at higher concentrations. Therefore, an additional experiment was performed to test their specificity and to rule out a property as a general protein precipitant. For this purpose, the human tumor suppressor p19^{INK4d} was purified as previously described (Baumgartner et al., 1998). Chalcone B-1 did not induce aggregation of p19^{INK4d} when applied under the same experimental conditions.

Chapter 4

In silico and NMR Identification of Inhibitors of the IGF-I and IGF-Binding Protein-5 Interaction

4.1 Introduction

Recently the crystal structure of the insulin-like growth factor-I (IGF-I) in complex with the N-terminal domain of the IGF-binding protein-5 (IGFBP-5) was determined (Zeslawski et al., 2001). Computer screening was then employed to find potential inhibitors of this interaction using the crystal coordinates. From the compounds suggested by *in silico* screens, successful binders were identified by NMR spectroscopic methods. NMR was also used to map their binding sites and calculate their binding affinities. Small molecular weight compounds (FMOC derivatives) bind to the IGF-I binding site on the IGFBP-5 with micromolar affinities, and thus serve as potential starting compounds for the design of more potent inhibitors and therapeutic agents for diseases that are associated with abnormal IGF-I regulation.

4.2 Biological context

The insulin-like growth factors (IGF-I and IGF-II, ca. 50% identity with insulin) are potent mitogens that promote cell proliferation and differentiation (Wetterau et al., 1999; Hwa et al., 1999). Most of the effects of IGF-1 (70 amino acids) are mediated by binding to the type I IGF receptor (IGF-IR), a heterotetramer that has tyrosine kinase activity. The level of free systemic IGF

is modulated by the extent of binding to IGF binding proteins (IGFBPs) (Jones & Clemmons, 1995; Martin, 1999). Signaling at the target organ is induced by proteolytic cleavage of IGFBP in the complex by kallikreins, cathepsins, and/or matrix metalloproteinases, which releases IGF from the fragmented IGFBP and enables binding of IGF to the receptor (Wetterau et al., 1999; Jones & Clemmons, 1995; Martin, 1999). The IGFBP family comprises six proteins (IGFBP-1 to 6) that bind to IGFs with high affinity and a group of IGFBP-related proteins (IGFBP-rP 1-9), which bind IGFs with lower affinity. The proteins are produced in all tissues, typically however with tissue specific relative amounts of the various IGFBPs (Hwa et al., 1999). A key conserved structural feature among the six IGFBPs is a high number of cysteines (16-20 cysteines), clustered at the N-terminus (12 cysteines) and also but to a lesser extent at the C-terminus. The proteins share a high degree of similarity in their primary protein structure (identities around 30-40%), with highest conservation at the N- and C-terminal regions. It has been shown that these regions participate in the high-affinity binding to IGFs (Baxter et al., 1992; Clemmons, 2001). Full length IGFBP-5 is a 29 kDa protein. It is expressed mainly in the kidney, and is found in substantial amounts in connective tissues. Unlike other IGFBPs, IGFBP-5 strongly binds to bone cells because of its high affinity for hydroxyapatite. IGFBPs regulate not only IGF action but appear also to mediate IGF-independent actions, including inhibition or enhancement of cell growth and induction of apoptosis. Recently, the presence of specific cell-surface IGFBP receptors were discovered. IGFBP-3 and IGFBP-5 have recently been shown also to be translocated into the nucleus, compatible with the presence of a nuclear localization sequence (NLS) in their mid-region. This raises the possibility that nuclear IGFBP may directly control gene expression (Baxter, 2001). IGFBPs were also shown to bind to important viral oncoproteins such as HPV oncoprotein E7 (Wetterau et al., 1999). The IGFs, with their potent mitogenic and antiapoptotic effects, have been widely studied for their role in cancer (Khandwala et al., 2000; Hankinson et al., 1998; Holly, 2000; Wolk, 2000). Serum IGF-I and IGFBP-3 have been proposed as candidate markers for early detection of some cancers. In addition, IGF-I and IGF-II exhibit neuroprotective effects in several forms of brain injury and neurodegenerative disease (Loddick et al., 1998). This implies that targeted release of IGF from their binding proteins in brain tissue, for example, might have therapeutic value for stroke and other neurodegenerative diseases (Loddick et al., 1998). Compounds which disrupt the IGFBP-IGF interaction thus represent potential drugs. This idea has been explored by Liu et al. (2001), who screened successfully a large library of compounds to identify molecules that could displace IGF from its binding proteins. In a structure based attempt to identify IGF releasing substances, the computer docking

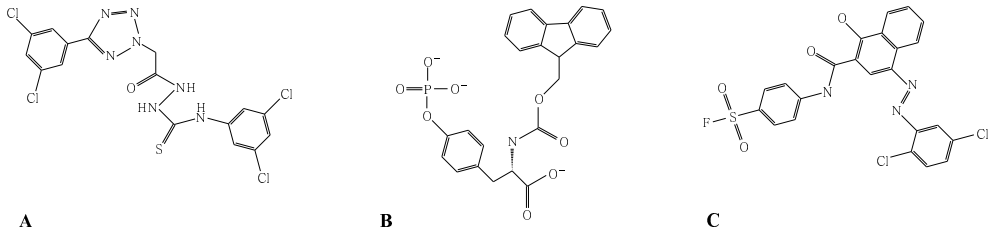

Figure 4.1: Formula of the compounds proposed by FlexX screening. (A) N1-(3,4-dichlorophenyl)-2-2-[5-(3,5-dichlorophenyl)-2H-1,2,3,4-tetraazol-2-yl]acetylhydrazine-1-carbothioamide (B) N-α-FMOC-O-phospho-L-tyrosine (C) 4-(2,5-dichlorophenylazo)-4'-fluorosulfonyl-1-hydroxy-2-naphthanilide.

program FlexX identified IGFBP-5 ligands, FMOC derivatives, that bind to the IGF-I binding site on IGFBP-5 with a micromolar affinity. These results should aid the search for more potent inhibitors of the IGF-I and IGFBP-5 interactions and thus potential IGF-I releasing therapeutics.

4.3 Results and Discussion

The FlexX program (Rarey et al., 1996a,b) and the crystal structure of the IGF-I complex with the N-terminal mini-IGFBP-5 fragment (Zeslawski et al., 2001) was used to identify potential inhibitors of the N-terminus-IGFBP-5/IGF-I interaction. Screening through the ACD database identified three dissimilar compounds (figure 4.1) with a theoretically predicted binding capacity to the IGFBP-5 region responsible for IGF-I interaction. Then NMR was applied to test for the predicted ligand-protein interactions (Shuker et al., 1996; McAlister et al., 1996). Titrations of the ^{15}N-labeled mini-IGFBP-5 with the potential inhibitors revealed no binding affinity for compounds A and C to mini-IGFBP-5. This is not unexpected and is a common drawback of *in silico* screenings as the produced possible binding modes do not necessarily reflect real ligand binding. For this reason hits from virtual screening must be verified by other methods. Compound B, however, clearly altered the ^{15}N-HSQC spectrum of the protein, indicating binding of this compound to mini-IGFBP-5 (figure 4.4). Compound B, because of its low solubility in water, was initially dissolved in DMSO. Titration of the protein with DMSO (e.g. lacking compound B) as a control was also performed. To investigate the influence of DMSO on the compound B binding to the protein, compound B dissolved in PBS buffer (at a lower concentration) was also titrated. Dissociation constants were estimated by monitoring several amino acid residues that display ligand induced changes in ^{15}N-^{1}H chemical shift (figures 4.2, 4.3 and 4.4). The values

Table 4.1: Dissociation constant calculations for compound B or DMSO binding to IGFBP-5 using data from distinct amino acid residues. Given errors are due to the fitting procedure.

residue	**ligand in DMSO**	**ligand in PBS**	**DMSO**
	K_D[mM]	**K_D[mM]**	**K_D[mM]**
Y50	1.58 ± 0.09	1.82 ± 0.95	648 ± 370
L73	1.31 ± 0.17	2.93 ± 1.41	541 ± 306
L81	2.78 ± 0.30	2.88 ± 1.18	610 ± 343
S85	1.38 ± 0.10	2.33 ± 0.94	650 ± 373
Y86	1.90 ± 0.17	1.72 ± 0.99	783 ± 498
R87	1.64 ± 0.12	2.36 ± 1.00	921 ± 662
K91	2.42 ± 0.18	2.12 ± 1.03	719 ± 434
average:	$\mathbf{1.9 \pm 0.5}$	$\mathbf{2.3 \pm 0.4}$	$\mathbf{700 \pm 100}$

of the dissociation constants for ligand B dissolved in DMSO and in PBS were similar (1.86 and 2.31 mM, respectively; Table 4.1 and Figure 4.4). These residues are concentrated mostly in a contiguous region of the three-dimensional structure of the mini-IGFBP-5 (figure 4.5 A) which comprise the binding site of IGF-I.

Dissociation constants for compound B and mini-IGFBP-5 interactions are significantly higher than the constants for interactions of the mini-IGFBP-5 with IGF-I, which are in the nanomolar range (Kalus et al., 1998). In the gel filtration studies compound B was not able to abolish the IGF-I/IGFBP-5 interactions (data not shown). Compound B was, however, used as a starting lead compound in search for higher affinity inhibitors for the IGF-I and IGFBP-5 interaction. Analysis of the IGFBP-5 residues involved in the compound B binding, as resolved by the present NMR study (Figures 4.2 and 4.5) and confirmed by molecular modeling predictions (figure 4.5 B), show that the binding region is in a similar location to that responsible for interactions with IGF-I (Zeslawski et al., 2001). It was tried to find derivatives of compound B with enhanced binding to IGFBP-5. Analogs of compound B are commercially available as they are commonly used in peptide synthesis. The binding surface between IGF-I and mini-IGFBP-5 appears mostly hydrophobic (Zeslawski et al., 2001), so first a compound B derivative Nα-FMOC-O-tert-butyl-L-tyrosine was tested, where the hydrophilic phosphate group of B is replaced by a similarly sized hydrophobic tert-butoxy group (compound B1). This substitution resulted in an increase

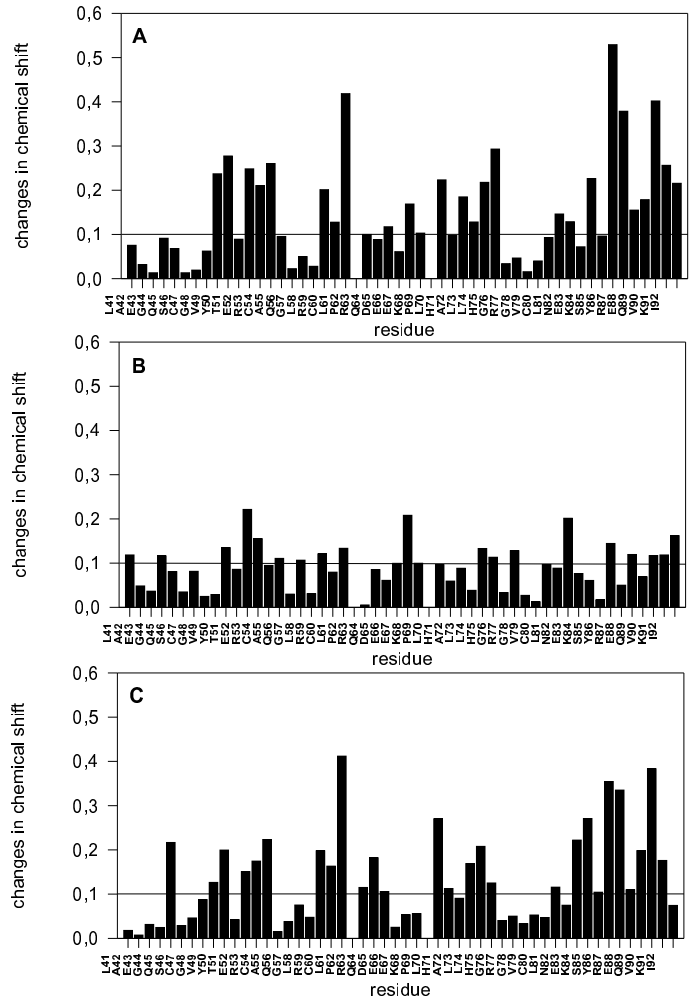

Figure 4.2: Differences in chemical shifts of free and inhibitor B-complexed mini-IGFBP-5 for all residues. Large shifts indicate residues involved in the compound B binding. Data for (A) ligand dissolved in DMSO (B) DMSO (C) ligand dissolved in PBS.

Table 4.2: Dissociation constants calculated for compound B and its derivatives binding to IGFBP-5 using changes in chemical shift for the residue L81.

compound	**chemical name**	$\mathbf{K_D[mM]}$
B	Nα-FMOC-O-phospho-L-tyrosine	2.78 ± 0.30
B1	Nα-FMOC-O-tert-butyl-L-tyrosine	0.718 ± 0.079
B2	Nα-FMOC-L-phenylalanine	1.075 ± 0.507
B3	Nα-FMOC-N-BOC-L-tryptophan	0.0432 ± 0.0115
B4	Nα-FMOC-L-leucine	1.088 ± 0.519

of the binding affinity by about threefold (table 4.2). The next compound tested resembled B1 but the tert-butyl group was completely omitted, resulting in Nα-FMOC-L-phenylalanine (compound B2). Binding of compound B2 was weaker than of compound B1 but still better than for compound B. The decrease in ligand binding affinity correlated with the reduction of compound size suggested that larger hydrophobic substituent may enhance affinity. Therefore an analog of compound B with a larger aromatic group (Nα-FMOC-N-BOC-L-tryptophan; compound B3) was tested; the substitution enhances ligand affinity into the micromolar range (43.2 μM; table 4.2). Substitution of the aromatic tryptophan by the aliphatic leucine did not improve the affinity of the binding (Nα-FMOC- L-leucine, compound B4, table 4.2). Compound B3, our best lead, was still not able to abolish IGF-I/IGFBP-5 interactions at concentrations tested in gel filtration studies (data not shown). Since it is well known that DMSO might have a considerable effect on proteins we finally performed two control experiments. Titration of the protein with DMSO (e.g. lacking compound B) as a control revealed very weak binding of DMSO to mini-IGFBP-5 (Figure 4.3 and table 4.1). The DMSO interaction is most likely non-specific, as indicated by the small and similar extent of the chemical shift perturbations of a large number of amino acid residues (Figure 4.3 B). Compound B was soluble in PBS buffer at low concentrations. Comparison of a titration of compound B in PBS and DMSO (Figures 4.3 A and B) shows that most significant changes appear at the same amino acid residues. Note that the changes in chemical shift do not necessarily go in the same direction for both experiments. So values in Figures 4.3 C and B might not be simply added to arrive at values in Figure 4.3 A, but will for different residues partially cancel or add up.

IGFs are known for their neuroprotective properties. Brain injury is commonly associated

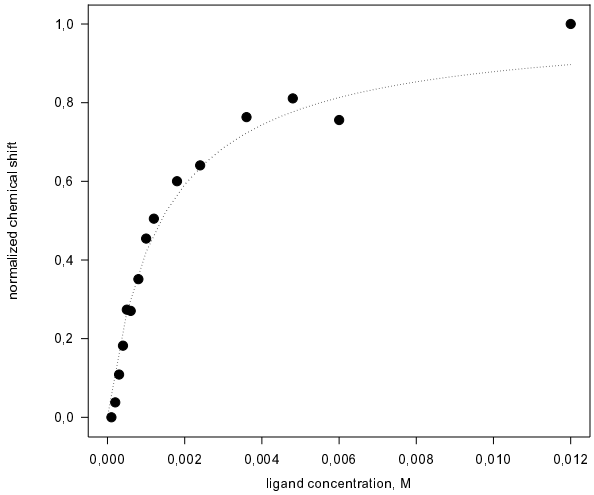

Figure 4.3: Titration of the mini-IGFBP-5 sample with the compound B dissolved in DMSO. Data for residue S85.

with increase in IGF expression but, paradoxically, also with increased expression of the inactivating binding proteins. Attempts to administer IGF-I exogenously as protective therapy in cases of brain injury (Gluckman et al., 1992) may thus be hampered by the increased expression of brain IGFBP. Combined administration of IGFs and IGFBP ligand inhibitors may optimize treatment of neurodegeneration. Alternatively, displacement of the large "pool" of endogenous IGF from the IGF-binding proteins might elevate "free" IGF levels such that administration of IGFBP ligand inhibitors elicit neuroprotective effects comparable to those produced by administration of exogenous IGF. Bayne et al. (1990) reported an IGFBP ligand inhibitor, [Leu24,59,60, Ala31] IGF-I mutant, with high affinity to IGF-binding proteins (0.3 - 3.9 nM) but with no biological activity at the IGF receptors ($> 10\mu$M). Loddick et al. (1998) examined effects of this high-affinity IGFBP ligand inhibitor in in vitro studies of release of "free" bioactive IGF-I from rat cerebrospinal fluid and in in vivo studies to evaluate its neuroprotective effects in a rat model of focal ischemia. This successful targeting of IGFBPs suggests that it may be possible to identify non-peptide small molecules that act as IGFBP ligand inhibitors, with the potential for good blood-brain barrier penetration and oral activity. The data collected by Loddick et al. (1998) demonstrate that displacement of IGFs from IGFBPs in the brain is a potential treatment for stroke. Moreover, in view of the potent actions of IGFs on survival of neurons and glial cells

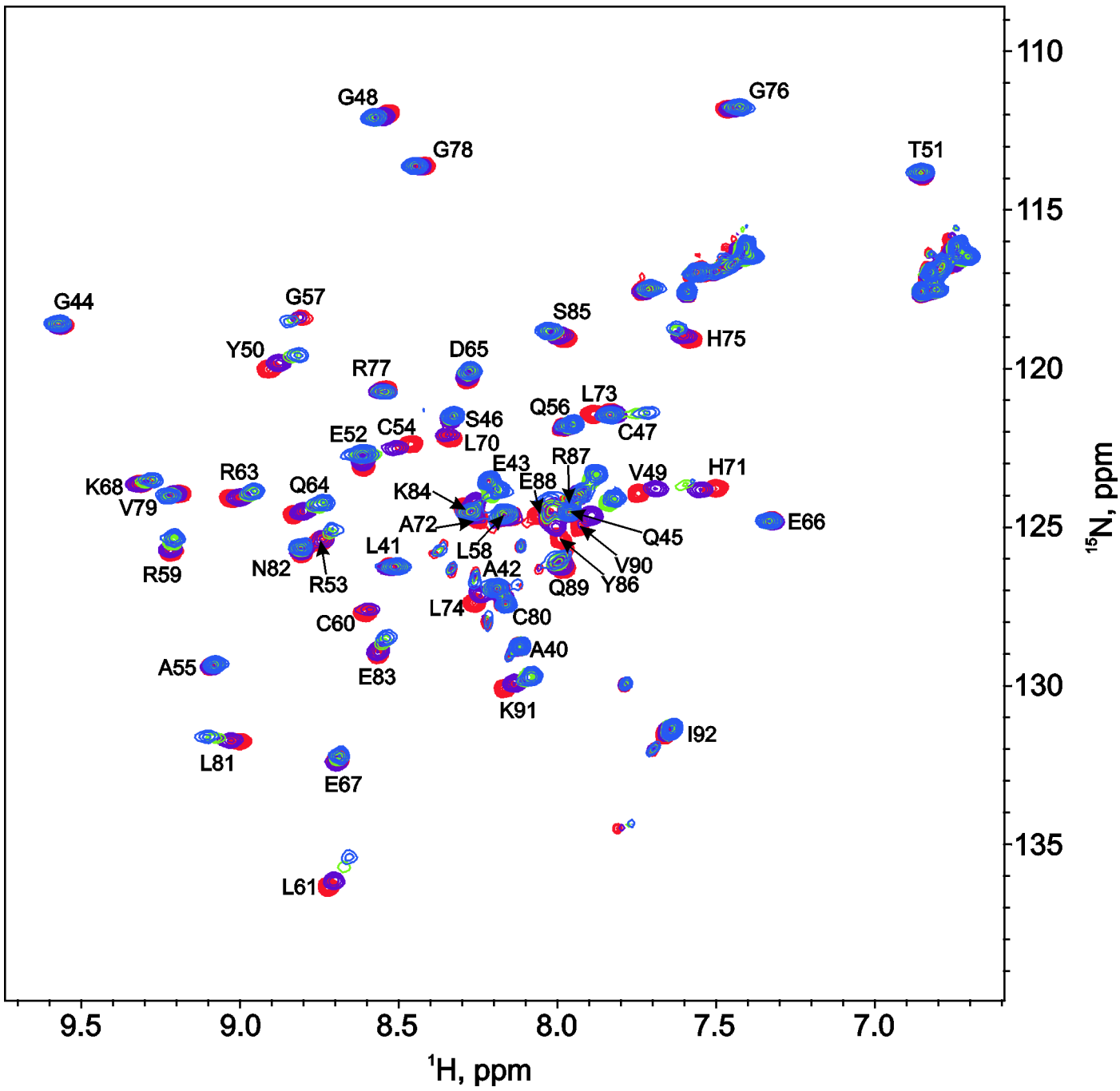

Figure 4.4: ^{15}N-HSQC spectrum illustrating the titration of the mini-IGFBP-5 with the increasing amounts of compound B. The reference is shown in red. 1:1, 1:5 and 1:10 titration steps (protein : ligand) in purple, green and blue, respectively.

as well as the widespread protective affects against a variety of brain insults, IGFBP ligand inhibitors may have broader utility for the treatment of various neurodegenerative disorders as well as traumatic brain and spinal cord injury.

Conclusion

Because of their high structure similarity, it was assumed that all B analogs bind similarly to IGFBP-5. This is supported by the fact, that mostly the same residues of IGFBP-5 are affected in the NMR titrations. Figure 4.5 shows compound B docked in the IGF-I binding site of IGFBP-5 and overlaid with IGF-I. Analysis of the structures shows the prediction that the phenyl group of the compound B mimics Phe16 from IGF-I (figure 4.5), and that the FMOC-group binds at the equivalent position of IGF-I-Leu54. The Glu3 binding region of IGFBP-5, however, seems not to be involved in interactions with compound B. Thus, this region offers binding interactions for new IGFBP-5 ligands, which when combined with compound B3 could significantly enhance binding affinities.

4.4 Experimental Section

Molecular Modeling

The protein model for flexible docking was taken from the high resolution X-ray structure of the IGF-I/mini-IGFBP-5 complex (Zeslawski et al., 2001) without further modification, i.e. the model neither underwent additional minimization nor were any side chain conformations changed. As the small molecule database, the Available Chemicals Directory (ACD, MDL Information System) of commercially available compounds was used and filtered to include approximately 90,000 compounds with Mr $\leq$ 550 Da that contain at least one atom from the set N, O, F, S. The stereo chemical information was used as provided by ACD. The set of molecule files were converted to the mol2 format with SYBYL (Tripos, St. Louis) with all hydrogens added. This set served as input to FlexX (GMD, St. Augustin) for flexible docking into a binding site on IGFBP-5 to identify small molecules which might bind to IGFBP-5 and thereby block the interaction with IGF-I. The binding site was defined as a sphere around all residues of IGFBP-5 towards the interaction site plus a 5 Å border (taking whole residues). The side chain conformations of mini-IGFBP-5 were not adjusted by the docking protocol. The small molecule conformations for each compound generated by FlexX using the standard FlexX scoring function were clustered

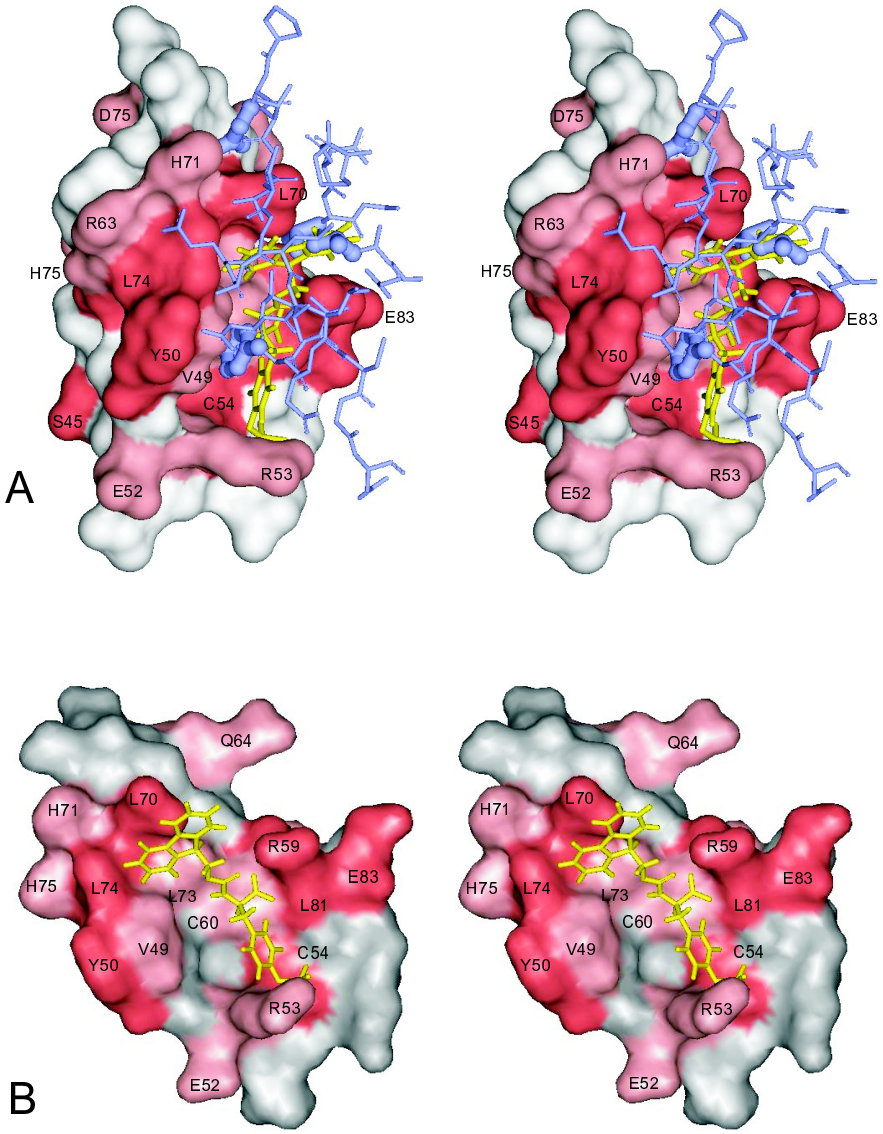

Figure 4.5: (A) Surface plot of mini-IGFBP-5 as resolved by X-ray crystallography superimposed with the docking result of compound B (yellow) and with the interface residues of the IGF-I/mini-IGFBP-5 complex. IGF is shown in blue. Four IGF-I residues most essential for interactions with IGFBP-5 (from the top: Glu3, Leu57, Leu54 and Phe16, respectively) are shown as blue balls. Residues with chemical shift changes due to binding of compound B as revealed by the present study are shown in red (the more intense the color the bigger changes). (B) A close-up of the mini-IGFBP-5 and compound B only.

by an r.m.s.d. of 2.3 Å and each best scoring pose within a cluster was saved as the cluster representative. Analysis of all the saved conformations of all docked ligands was carried out using a distance-based filter defining the following criteria: (1.) A substructure of the ligand must interact with the region Val49/Leu70/Leu73/Leu74. (2.) A substructure of the ligand must interact in the deep pocket around Cys47/Thr51. As a result, three compounds were selected for an NMR screening (Figure 4.1).

Materials

Mini-IGFBP-5 (amino acids 40-92 of human IGFBP-5) was expressed and purified using the construct described by Kalus et al. (1998). Compounds A, B and C were purchased from ChemPur (Karlsruhe, FRG), Fluka (Buchs, Switzerland) and Sigma (Deisenhofen, FRG), respectively. Compound B derivatives were generously provided by Prof. Luis Moroder.

NMR assignment

Previously the NMR assignment of IGFBP-5 has been reported by Kalus et al. (1998) at pH 4.7. The ligand binding studies reported here were performed at a more physiological pH value of 7.2. Even after several pH titration experiments, the assignment of the amide groups in the HSQC spectra could not be transfered completely. Several important residues could not be traced through all titration steps. To resolve the assignment, a 2-D NOESY and a ^{15}N-NOESY-HSQC spectrum were recorded. Thus NOESY patterns from each residue could be compared to those assigned by Kalus et al. (1998). From the structure (available under the PDB ID: 1BOE at the Brookhaven Data Bank, www.rcsb.org/pdb; Berman et al. (2000)) distance constraints were also used to identify NOESY crosspeaks and thus backbone amide groups in the 2-D and 3-D NOESY spectra (see Figure 4.6). Additionally a selectively ^{15}N leucine labeled sample was prepared to verify the assignment of the crucial residues Leu70, Leu73 and Leu74. The HSQC spectrum from this selectively ^{15}N leucine labled sample superimposed on the HSQC of the uniformly labeled sample is shown in Figure 4.7. The single isoleucine present in this protein shows a peak as strong as those from the leucines due to cross-labeling. Also all three valins can be identified, their resonances slightly less intense. Interestingly, leucine 74, which is involved in ligand binding could not be assigned unambigously from this spectrum. It could be identified in the 2-D NOESY though (Figure 4.6). The corresponding peak is also indicated in the HSQC in Figure 4.7. For the complete assignment see Figure 4.4.

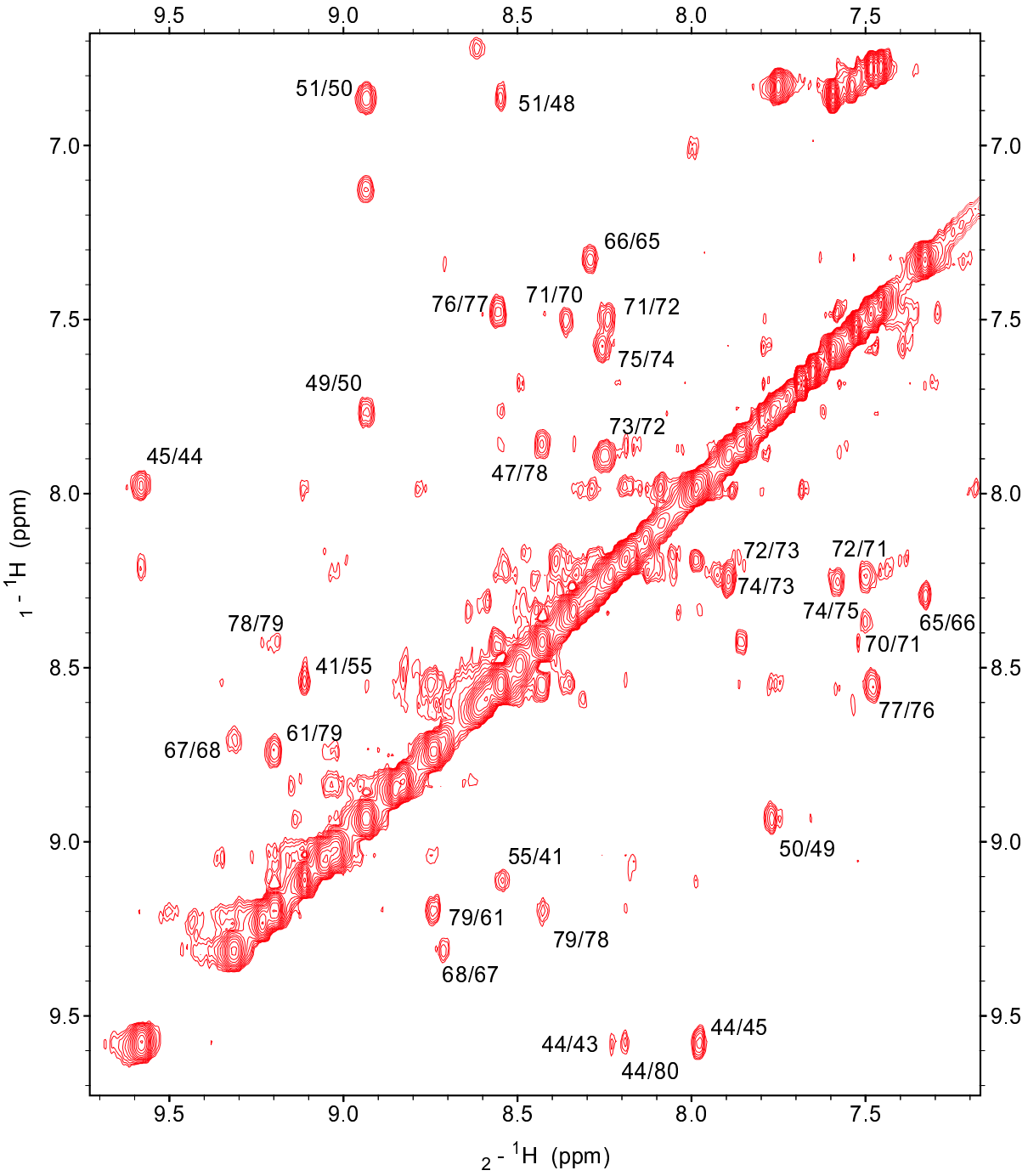

Figure 4.6: Part of the 2D NOESY spectrum of IGFBP-5. Some of the cross-peaks which were crucial for the assignment are labeled with their respective sequence numbers.

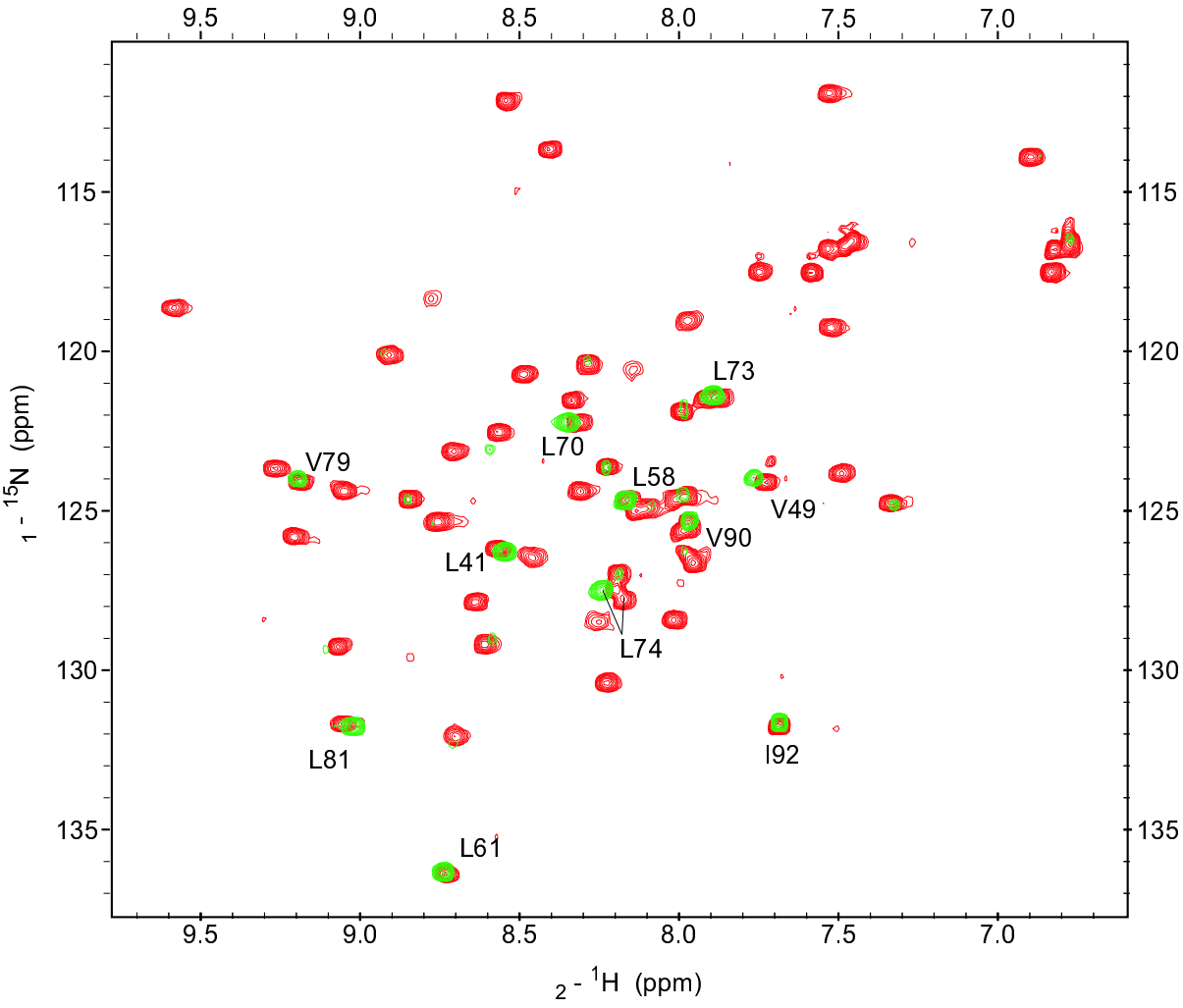

Figure 4.7: The HSQC spectrum from the selectively ^{15}N leucine labeled sample superimposed on the HSQC of the uniformly labeled sample. The assignment is given only for leucine, valine and isoleucine residues. Leucine 74 was ambiguous.

Detection of Ligand Binding

Ligand binding was detected by acquiring ^{15}N-HSQC spectra. All NMR spectra were acquired at 300 K on Bruker DRX600 spectrometer. The samples for NMR spectroscopy were concentrated and dialyzed against PBS buffer. Typically, the sample concentration was varied from 0.3 to 1.0 mM. Before measuring, the sample was centrifuged in order to sediment aggregates and other macroscopic particles. 450 μl of the protein solution were mixed with 50 μl of D_2O (5-10%) and transferred to an NMR sample tube. The stock solutions of compounds were 100 mM either in water or in perdeuterated DMSO. pH was maintained constant during the whole titration. The binding was monitored by observation of the changes in the ^{15}N-HSQC spectrum. Dissociation constants were obtained by monitoring the chemical shift changes of the backbone amide of several amino acid residues (Table 4.1) as a function of ligand concentration. Data were fit using a single binding site model.

Chapter 5

A novel Medium for Expression of selectively ^{15}N labeled Proteins in SF9 insect cells

5.1 Introduction

In the last years a growing number of proteins was expressed using the baculovirus expression vector system. Whereas bacterial expression systems are widely used for production of uniformly or selectively labeled proteins the usage of the baculovirus expression system for selective labeling is limited to very few examples in the literature. Two insect media, IML406 and IML455 for the production of selectively labeled protein in insect cells were recently developed in our group. The same levels of cell densities and proteins compared to other insect media could be obtained. The utilized amounts of ^{15}N-amino acids for the production of labeled GST as a sample protein were similar in the case of bacterial and viral expression. For most amino acids the ^{15}N-HSQC spectra, recorded with GST labeled in insect cells, showed no cross-labeling and provided therefore spectra of better quality as compared to the NMR spectra of the protein expressed in *E. coli.* The reason was the large number of amino acids, which are essential for insect cells. Also in the case of non-essential amino acids selective labeling could be accomplished. Therefore the selectively labeling using the baculovirus expression vector system represents a complement or even a powerful alternative to the bacterial expression system. The quality of the new media and the extent of cross-labeling in the baculovirus system was monitored by ^{15}N-HSQC experiments.

5.2 Biological context

The baculovirus based expression systems are among the most powerful expression systems known in biochemistry. They have several advantages over bacterial expression systems, as they allow for simple production of functional heterologous proteins like, for example, enzymes (Lawrie et al., 1995; Kumar et al., 2001), antibodies (Brocks et al., 1997) and receptors (Cascio, 1995; Zhu et al., 2001). There are also a number of proteins that can only be efficiently expressed in their folded and functional forms in insect cells. High cell densities are needed for obtaining high protein yields and therefore several media (Doverskog et al., 1998; Ferrance et al., 1993) and feeding strategies (Kim et al., 2000; Doverskog et al., 2000; Mendonça et al., 1999; Chiou et al., 2000) were developed to enhance cell growth. Identification of essential and non-essential amino acids for cell growth is also important. So far, alanine, cysteine, glutamic acid, glutamine, aspartic acid and asparagine were found to be non-essential amino acids (Öhman et al., 1996; Doverskog et al., 1998). The other amino acids are supposed to be essential for insect cells. The insect cells require also growth factors, vitamins and other compounds for higher cell densities (Mendonça et al., 1999; Öhman et al., 1995). These components are provided by chemically-not-defined substances, like yeastolate or fetal calf serum (Drews et al., 1995; Ferrance et al., 1993).

NMR-based structural studies and NMR-based ligand binding studies require selectively and/or uniformly ^{15}N-labeled proteins. For expression of uniformly labeled proteins in *E. coil* ^{15}N-ammonium chloride can be used as a sole source for nitrogen, whereas commercially available media for uniformly labeling in insect cells contain all ^{15}N-amino acids, which increase the costs dramatically. For selectively labeling in bacteria a medium is used that contains all amino acids in a similar manner as for insect cells. In this case comparable costs can be expected. Only few reports can be found in the literature on labeling proteins in insect cells (Creemers et al., 1999; DeLange et al., 1998), in addition the total composition of the used media was kept confidential. A novel optimized medium to label proteins expressed in Sf9 insect cells was developed in our laboratory, using the glutathione-S-transferase protein (GST) as a model. This protein was chosen because it expresses well in *E. coli* and therefore labeling can be compared with that in SF9 insect cells. It also possesses high stability and has a size of 27 kDa, which is in a typical range for many proteins expressed with the baculovirus expression vector system. The goal of the work was to develop a novel medium and to investigate the possibility of uniformly and selectively labeling proteins with ^{15}N-amino acids in insect cells.

5.3 NMR Spectroscopy

To investigate the quality of the new media IML455 and IML406, selectively ^{15}N labeled samples of GST were expressed using ^{15}N- glycine, leucine, lysine, phenylalanine and valine, which are in general easily labeled in bacterial expression systems. In a further step aspartic acid, glutamic acid and ammonium chloride were used for labeling studies in Sf9. The formulation of the novel media is reported by Brüggert (2002). The medium IML455 contained NH_4Cl instead of aspartic acid and glutamine used in IML406. To assess the quality of the new media, also selectively labeled samples from the *E. coli.* system were prepared. The bacterial media for selective labeling of proteins was prepared as described (Senn et al., 1987). Just after induction the same amount of the ^{15}N-labeled amino acid as that used in the medium was added.

For the NMR experiments the protein solutions were concentrated with a Centricon10 (Amicon) to the volume of 450 ml and 50 ml D_2O (99.9%) was added to the sample. The sample concentration ranged from 0.2 to 0.8 mM. All NMR spectra were acquired at 300 K on a Bruker DRX-600 spectrometer. ^{1}H-^{15}N-HSQC spectra (Mori et al., 1995) were recorded with 128 increments in the indirect ^{15}N dimension with a number of scans varying from 4 to 1024 depending on the concentration of individual samples. Measurement times ranged thus from 2 to 24 hours. Processing and analysis of the spectra was performed using the programs xwinnmr (Bruker) and Sparky (Goddard & Kneller, 2001), respectively.

5.4 Results and Discussion

For protein-ligand binding studies or for structure determination with NMR on proteins expressed in insect cells the use of selectively labeled protein samples is essential. Whereas uniformly labeled media are available from several companies the possibility to obtain selectively labeled media is restricted. This kind of medium is only prepared on request and the composition is secret. The formulation of a medium, which can be flexibly utilized for selectively labeling, is highly desired. This medium has to fulfill requirements for high level expression with as small as possible amount of amino acids. The media IML 406 and IML 455, developed in our laboratory, were used for ^{15}N-labeling studies using ^{1}H-^{15}N-HSQC experiments.

Table 5.1: Used amounts of ^{15}N-compounds, number of peaks visible in the ^{1}H-^{15}N-HSQC spectra (the expected number is given in parenthesize as the C-terminus in the proteins from the two expression systems is not identical) and the number of corresponding peaks, each given for expression in *E. coli.* or Sf9 cells.

	Amount of ^{15}N-compound used in medium		**Number of signals (strong/weak(expected))**		**identical signals (strong/weak)**
^{15}N-compound	***E. coli***	**Sf9**	***E. coli***	**Sf9**	
GLY	800 mg/l	650 mg/l	16/-(17)	17/-(16)	15/-
LYS	625 mg/l	200 mg/l	19/-(21)	19/2(21)	17/-
VAL	200 mg/l	200 mg/l	13/7(10)	11/6(10)	10/-
PHE	100 mg/l	250 mg/l	18/6(9)	9/3(9)	8/2
LEU	200 mg/l	400 mg/l	27/4(28)	34/10(28)	22/-
GLU	800 mg/l	600 mg/l	50/18(16)	40/4(16)	31/-
ASP	500 mg/l	350 mg/l	49/18(18)	0/0(19)	-

^{15}N-labeling with glycine, lysine, valine, phenylalanine and leucine

For the labeling studies GST was expressed in Sf9 and *E. coli* using single ^{15}N-amino acids. Table 5.1 gives an overview of the amounts of amino acids used for different media. The selective labeling with ^{15}N-amino acids in Sf9 results in most cases in equal or better quality NMR spectra compared to the spectra obtained from GST expressed in bacteria. Since most of the amino acids are essential for insect cells a conversion between these amino acids is not possible. The labeling with ^{15}N-glycine and ^{15}N-lysine lead to nearly identical results in both expression systems. The positions of most of the peaks in the spectra were identical. The number of signals was close to the number of both amino acids in GST (see Table 5.1). In insect cells no cross-labeling from glycine to the essential amino acids serine is detectable. In *E. coli* this conversion is suppressed by an excess of ^{14}N-serine in the medium. In the case of ^{15}N-lysine in both expression systems no cross-labeling is visible. This amino acid is not efficiently used for formation of other amino acids. Two peaks differ in chemical shift depending on host cell indicating pH-sensitivity or posttranslational modifications of these lysines.

The problem of cross-labeling in *E. coli* appeared for labeling with ^{15}N-phenylalanine and

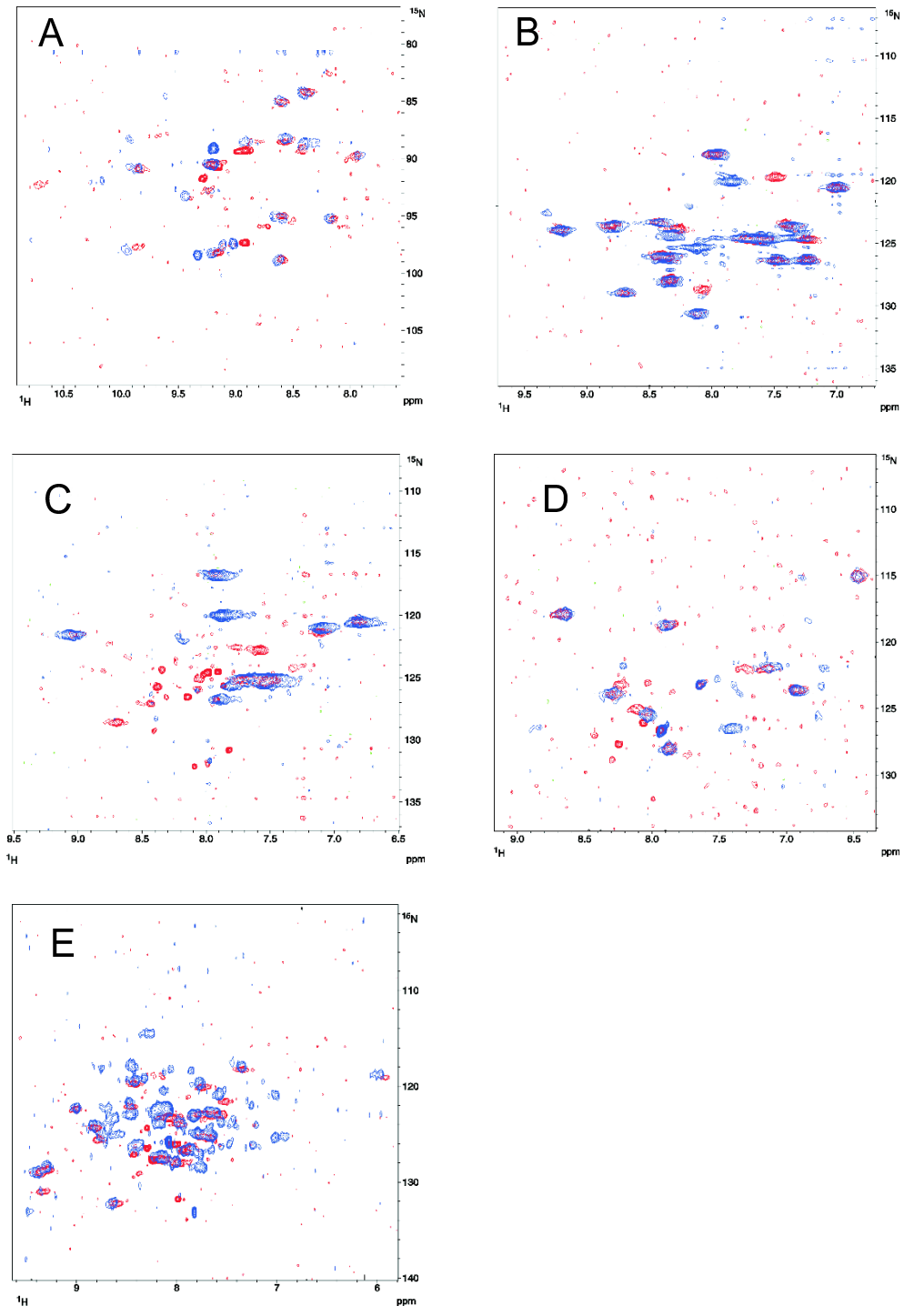

Figure 5.1: ^{15}N-labeling with glycine, lysine, valine, phenylalanine and leucine. Superposition of HSQC spectra of selectively labeled GST from Sf9 (blue) and *E. coli* (red) A) ^{15}N-glycin, B) ^{15}N-lysin, C) ^{15}N-phenylalanine, D) ^{15}N-valine and E) ^{15}N-leucine.

^{15}N-valine. In the HSQC-spectra of GST expressed in bacteria the number of signals clearly exceed the number of the two amino acids. Phenylalanine is converted to tyrosine and valine to alanine. These reactions are not detectable in Sf9. Tyrosine is essential for insect cells and cannot be formed from phenylalanine. The formation of alanine from valine is also not visible, though the medium IML406 contains no additional alanine in contrast to the medium for bacterial expression. This pathway seems not to be efficient in Sf9. The use of ^{15}N-leucine in IML406 yields a HSQC-spectrum in which the number of intensive signals clearly exceeds the number of leucines in GST. This indicates cross-labeling via a yet unknown pathway. In the case of bacterial expression the number of intensive peaks is lower and close to the number of leucines in GST. In *E. coli* leucine is used for the formation of isoleucine. Isoleucine is converted in a second step to valine. Both reactions are not possible in Sf9. In conclusion, for ^{15}N-labeling with glycine, lysine, valine, phenylalanine or leucine cross-labeling is mainly a problem of bacterial expression.

^{15}N-labeling with ammonium chloride, glutamic acid or aspartic acid

In a further step the possibility to use ammonium chloride, aspartic acid or glutamic acid for ^{15}N-labeling of GST was investigated. It was reported that ammonium is used for formation of alanine and the amide group of glutamine (Drews et al., 2000). This would be a cheap alternative for selective labeling for alanine. The incorporation of $^{15}NH_4$ in the amide groups of arginine and glutamine could be confirmed (Figure 5.2 A). The formation of ^{15}N-alanine was not detected. The reason may be that the efficient formation of ^{15}N-alanine from $^{15}NH_4$ starts 72-80 hours after the beginning of fermentation (Drews et al., 2000). At this point in time the infected cells are already harvested. In both expression systems a powerful conversion between glutamic acid and aspartic acid was detected (Figure 5.2 B). In insect cells no spectrum for GST labeled with ^{15}N-aspartic acid could be obtained. The reason is the high amount of unlabeled asparagine, glutamine and glutamic acid, which is about six times higher than the content of ^{15}N-aspartic acid in IML406. The labeling of ^{15}N-aspartic acid is efficiently suppressed. In *E. coli* the fraction of ^{15}N-aspartic acid is higher resulting in an interpretable spectrum (Figure 5.2 B). For efficient labeling in insect cells the amount of this amino acid has to be increased. A better alternative is to remove aspartic acid from the medium and to label this amino acid and glutamic acid simultaneously. The higher amount of ^{15}N-glutamic acid in IML 406 yields a spectrum of good quality (Figure 5.2 B). The formation of alanine from glutamic acid is better visible

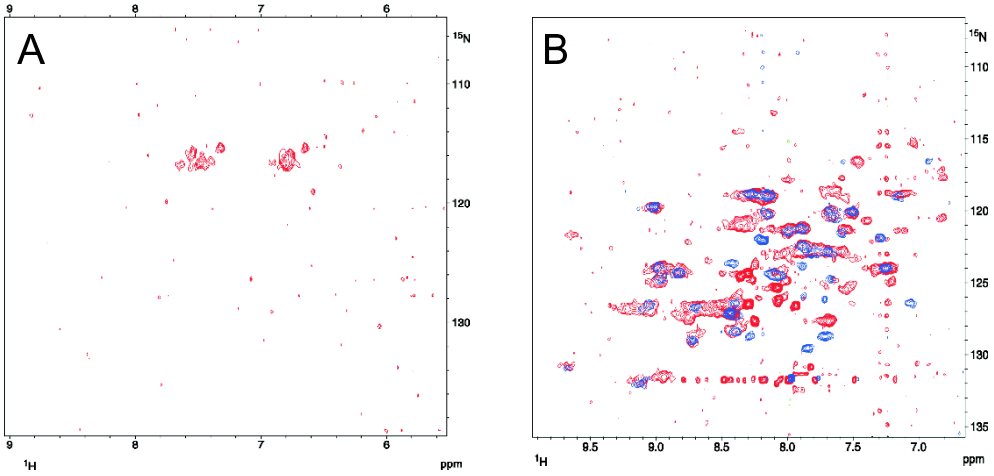

Figure 5.2: ^{15}N-labeling with ammonium chloride, glutamic acid or aspartic acid. A) HSQC spectrum of GST labeled with ^{15}N-ammonium chloride in SF9 cells. B) Superposition of the HSQC spectra of GST clabeled with ^{15}N-glutamic acid in SF9 (blue) and ^{15}N-aspartic acid in *E. coli* (red)

in Sf9 than in bacteria. In contrast to the medium for bacterial expression IML406 contains no additional alanine and the cross-labeling is not suppressed. The total number of peaks in the spectrum of GST labeled with ^{15}N-glutamic acid in Sf9 was only lower by four peaks than the total number of aspartic acid, glutamic acid and alanine in GST. Thus these three amino acids can be labeled efficiently using ^{15}N-glutamic acid. No cross-labeling to other amino acids is visible. *E. coli* use glutamic acid additionally to form many other amino acids like valine, phenylalanine, leucine or tyrosine. The transamination between these amino acids plays a central role in bacteria. In Sf9 this reaction is limited to amino acids involved in the TCA cycle and alanine.

Using $^{15}NH_4Cl$ and ^{15}N-glutamic acid, glutamic acid, glutamine, aspartic acid, asparagine and alanine can be labeled simultaneously in IML406. Based on these findings a first simplified overview of the network of the amino acid metabolism in *E. coli* and insect cells focused on nitrogen may be presented. Figure 5.3 shows that the transamination is limited to a few reactions, whereas in *E. coli* the nitrogen is transferred between most of the amino acids. Especially the central role of glutamic acid for synthesis of amino acids is clearly visible.

Figure 5.3 also shows that the expression in Sf9 is better suited for selectively labeling of amino acids. The expression system offers the possibility to selectively label amino acids in high degrees and without cross-labeling including tyrosine, phenylalanine, glycine, serine, cysteine, arginine and valine. Even amino acids can be labeled, which are normally not used in *E. coli*

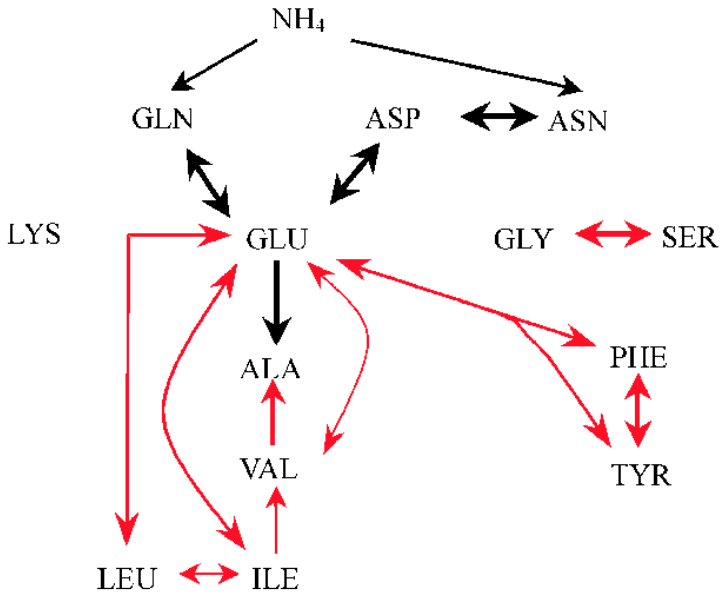

Figure 5.3: Simplified presentation of the amino acid metabolism in *E. coli* and Sf9 with respect to ^{15}N: Pathways which are present in both organisms are shown in black; pathways which exist only in *E. coli* are depicted in red.

due to extensive cross-labeling. The expression in insect cells is in many cases a potential alternative to the in-vitro-expression. This is important for proteins, which are not available in high yields or in a functional form in bacteria. In principle the labeling in insect cells offers the opportunity for the NMR structure determination of proteins expressed in Sf9. In addition the possibility of selectively labeling using all essential amino acids accelerates the assignment of signals for proteins which are expressible in bacteria and insect cells in comparable yields.

Chapter 6

NMR Characterization of the cyclase-associated protein (CAP) from *Dictyostelium discoideum*

6.1 Biological context

The cyclase-associated protein (CAP) of *Dictyostelium discoideum* is an actin binding protein that is involved in the microfilament reorganization at anterior and posterior plasma membrane regions (Gottwald et al., 1996). CAP was first isolated from *Saccharomyces cerevisiae* as a component of the adenylyl cyclase (Cyr1p) complex (Field et al., 1990) and the protein is believed to act as one of the bridging proteins that link nutritional response signaling and changes in the actin cytoskeleton (for a review see Hubberstey & Mottillio (2002)). *Dictyostelium* CAP is a bifunctional protein as well. While the actin sequestering activity has been localized to the carboxy-terminal 210 amino acids, the N-terminus seems to mediate this activity in a PIP_2-regulated manner (Gottwald et al., 1996). The amino-terminal domain has also been shown to localize the whole protein to the membrane (Noegel et al., 1999).

6.2 The folded core of CAP-N

In general *Dictyostelium* CAP follows the domain organization of all CAP-homologues as given by Gerst et al. (1992): it consists of an amino-terminal domain encompassing residues 1-215 and a carboxy-terminal domain encompassing residues 255-464, separated by a proline-rich

linker domain of 39 residues. Our NMR characterization of an amino-terminal construct (CAP-N) encompassing residues 1-226 revealed a different domain structure though (see also section 2.3).

Attempts to crystallize this construct failed which is frequently an indication for unstructured and highly flexible regions in the protein. After the sample was left at room temperature for one week, a one-dimensional proton NMR spectrum showed degraded peptide fragments (Figure 2.4 B). The linewidth of the protein signals did not improve as compared to the initial spectrum taken of the freshly prepared sample (Figure 2.4 A). It was assumed, that the sample contained now a mixture of several protein fragments of different length as well as cleaved off peptide fragments. Mass spectrometry confirmed the presence of protein fragments ranging in length from 226 to 173 amino acids. Edmann sequencing showed, that all fragments had been cleaved at their amino terminus.

Further clues to the exact length of the folded core of the protein were given by a number of selectively ^{15}N-labeled samples. As an example Figure 6.1 shows two HSQC spectra from a selectively ^{15}N-alanin labeled sample superimposed on the spectrum of the uniformly labeled short construct that was later used for our NMR investigation (see below). The assignment for all alanin residues visible in this part of the spectrum is given according to their position in the sequence of this short construct i.e. from 1 to 176 corresponding to residues 51-226 of the original construct. The freshly prepared ^{15}N-alanin sample (residues 1-226) is shown in black. The sample was then left for seven days at room temperature and cleaved peptide fragments removed by dyalisis. Then another spectrum was recorded, here shown in green. Seven peaks have either completely disappeared or are now very weak (one of them not so clearly visible under the peak labeled A40). This is a clear indication, that the part of the sequence, which is cleaved off should include seven Alanin residues. Note, that the disappeared peaks all lie at a proton frequency of about 8.3 ppm (the random coil value, see also chapter 2) and thus represent an unfolded part of the protein.

Together with similar information from other selectively labeled samples it could be concluded, that about 50 amino acid residues are cleaved off the N-terminus of the protein. In this case it was very favorable that all together eight selectively labeled samples were available as the interpretation of the spectra is not necessarily straight forward. In the case shown here residue A175, which is very close to the C-terminus of the fragment, does not show up in the selectively labeled sample. Also the rather strong peak in the center of the spectrum (8.2/124.5 ppm) later in the assignment turned out to be lysin 75. The peak at 8.4/126 ppm represents an

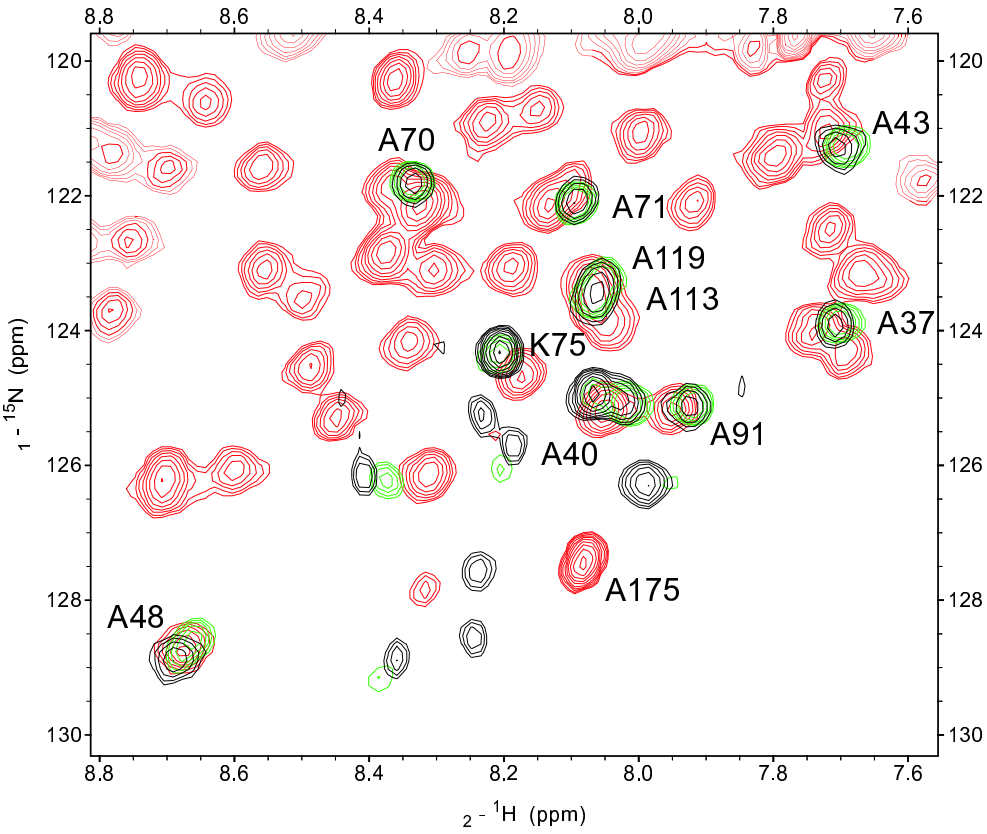

Figure 6.1: Part of ^{1}H-^{15}N-HSQC spectra of CAP-N from *Dictyostelium discoideum*. Superposition of spectra from an uniformly ^{15}N-labeled sample of the short construct (red), the selectively labeled ^{15}N-alanin sample of the long (black) and short (green) construct. The assignment for the alanin residues and for lysin 75 is indicated. For details see text.

Table 6.1: Sequence of the final, 176 residue construct of CAP-N

1	SVKEFQNLVD	QHITPFVALS	KKLAPEVGNQ	VEQLVKAIDA	EKALINTASQ
51	SKKPSQETLL	ELIKPLNNFA	AEVGKIRDSN	RSSKFFNNLS	AISESIGFLS
101	WVVVEPTPGP	HVAEMRGSAE	FYTNRILKEF	KGVNQDQVDW	VSNYVNFLKD
151	LEKYIKQYHT	TGLTWNPKGG	DAKSAT		

A (ALA)	13	G (GLY)	9	M (MET)	1	S (SER)	15
C (CYS)	0	H (HIS)	3	N (ASN)	13	T (THR)	9
D (ASP)	7	I (ILE)	9	P (PRO)	8	V (VAL)	15
E (GLU)	13	K (LYS)	17	Q (GLN)	9	Y (TYR)	4
F (PHE)	9	L (LEU)	15	R (ARG)	4	W (TRP)	3

eighth alanin which is no longer present in the short construct (red peaks) as there are actually eight alanins among the first 50 residues.

A new fragment was subsequently cloned and expressed in *E. coli*, which encompasses the 176 residues from positions 51-226 of the original construct. This shorter construct was not degraded even after several month as was proved again by one-dimensional proton NMR. This sample was successfully crystallized (Ksiazek, 2002) and subjected to further NMR investigations.

In conclusion, NMR spectroscopy supplemented with mass-spectrometry and sequencing proved the structured amino-terminal domain of *Dictyostelium* CAP to exclude a serine-rich stretch at the N-terminus and to encompass the 176 residues from positions 51-226 (see Table 6.1).

6.3 Material and Methods

The cDNAs encoding the amino-terminal 226 residues of CAP plus a C-terminal His-tag (CAP-N'Px) or the 176 residues from position 51-226 (CAP-N) were cloned into the *Nde*I and *Bam*HI restriction sites of the pT7-7 expression vector (Tabor, 1990). *E. coli* BL21 harboring the plasmids were grown at 30°C (CAP-N'Px) resp. 37°C to an OD_{600} of 0.6-0.8. For expression of the

protein IPTG was added to a final concentration of 0.5 mM and cells were further incubated over night. After lysis and centrifugation in both cases the 100.000 $\times$ g supernatants were then purified on DE52 (Whatman) anion exchange, phosphocellulose (P11, Whatman) cation exchange, hydroxyapatite (Bio-Rad) and Ni-NTA (Qiagen) columns following standard procedures. The samples were finally concentrated in a Centriprep-10 concentrator (Amicon).

Uniformly ^{15}N-^{13}C and ^{15}N isotopically enriched protein samples were prepared by growing the bacteria in minimal media containing $^{15}NH_4Cl$, either with or without ^{13}C-glucose, respectively. For selectively enriched samples, defined media (Senn et al., 1987) were used that contained 100 to 800 mg/l of the isotopically enriched amino acids and all other amino acids.

The following samples were available in concentrations ranging from 0.8 to 1.2 mM at pH 7.3: Uniformly ^{15}N labeled CAP, as well as selectively ^{15}N-Ala, ^{15}N-Phe, ^{15}N-Gly, ^{15}N-Ile, ^{15}N-Lys, ^{15}N-Leu, ^{15}N-Val and ^{15}N-Gly/^{15}N-Ser labeled samples and a ^{15}N-^{13}C double labeled sample. All samples contained 10% D_2O. NMR spectra were recorded on Bruker DRX 600 and DMX 750 spectrometers equipped with triple resonance probeheads and pulsed-filed gradient units. The backbone resonances were assigned using a pair of HNCA and CBCA(CO)NH triple-resonance spectra with the help of ^{15}N-HSQC spectra recorded from the selectively labeled samples. Furthermore a HNCO and two 3D ^{15}N-NOESY-HSQC spectra with mixing times of $\tau_m = 120\text{ms}$ and $\tau_m = 40\text{ms}$ and a ^{13}C-NOESY-HSQC with a mixing time of $\tau_m = 100\text{ms}$ were used. All spectra were recorded at a temperature of 300 K. For a complete list of the recorded spectra see Table 6.2. Assignment was accomplished using the software package *sparky* (Goddard & Kneller, 2001).

6.4 Sequence-specific (^{1}H, ^{15}N, ^{13}C) resonance assignment

Figure 6.4 shows the quality and degree of overlap in the ^{1}H-^{15}N-HSQC spectrum. The signal dispersion in both the ^{15}N and ^{1}H dimensions is quite good, still assignment was complicated by considerable overlap in the central part of the spectrum. Superposition of this spectrum with HSQC spectra from the selectively labeled samples (see Table 6.2) lead to the unambiguouse identification of 6 out of 13 alanins, 8 of 9 phenylalanins, 7 of 9 glycins, 5 of 9 isoleucines, 8 of 15 leucines, 7 of 15 serins and 10 of 15 valins. Great caution had to be taken as the selectively labeled samples were of the longer construct of CAP-N while the uniformly labeled sample was of the short construct (see section 6.2). Starting from these identified residues, the CBCA(CO)NH and HNCA spectra were used to establish the sequence specific assignment.

Table 6.2: Recorded experiments and acquisition parameters

Experiment	spectral width $[\mathrm{ppm}]$			data points			scans
NOESY[a]	^{1}H:14.5	^{1}H:14.5	-	^{1}H:2048	^{1}H:790	-	256
TOCSY[b]	^{1}H:14.5	^{1}H:14.5	-	^{1}H:4096	^{1}H:600	-	256
^{1}H-^{15}N-HSQC[c]	^{1}H:11.6	^{15}N:29.6	-	^{1}H:2048	^{1}H:128	-	128
^{1}H-^{13}C-HSQC	^{1}H:17.4	^{13}C:74.6	-	^{1}H:2048	^{1}H:251	-	256
^{15}N-NOESY-HSQC[d]	^{1}H:16.0	^{15}N:35.1	^{1}H:16.0	^{1}H:2048	^{15}N:200	^{1}H:68	16(8[e])
^{13}C-NOESY-HSQC[f]	^{1}H:11.9	^{13}C:116.6	^{1}H:11.9	^{1}H:1024	^{13}C:160	^{1}H:160	16
CBCA(CO)NH	^{1}H:13.3	^{15}N:41.1	^{13}C:82.8	^{1}H:2048	^{15}N:64	^{13}C:64	96
HNCA	^{1}H:13.3	^{15}N:41.1	^{13}C:33.1	^{1}H:2048	^{15}N:64	^{13}C:64	32
HNCO	^{1}H:11.9	^{15}N:41.1	^{13}C:33.1	^{1}H:2048	^{15}N:64	^{13}C:64	32
HNHA	^{1}H:11.6	^{15}N:41.1	^{1}H:11.6	^{1}H:2048	^{15}N:128	^{1}H:38	32

[a] 2D-NOESY spectra were recorded with mixing times of $\tau_m = 40\mathrm{ms}$ and $\tau_m = 120\mathrm{ms}$ and on a sample in D_2O with $\tau_m = 20\mathrm{ms}$.

[b] $\tau_m = 30\mathrm{ms}$

[c] ^{1}H-^{15}N-HSQC spectra were recorded of uniformly and selectively ^{15}N-Ala, ^{15}N-Phe, ^{15}N-Gly, ^{15}N-Ile, ^{15}N-Lys, ^{15}N-Leu, ^{15}N-Val and ^{15}N-Gly/^{15}N-Ser labeled samples.

[d] ^{15}N-NOESY-HSQC spectra were recorded with mixing times of $\tau_m = 120\mathrm{ms}$ and $\tau_m = 40\mathrm{ms}$

[e] a ^{15}N-NOESY-HSQC ($\tau_m = 100\mathrm{ms}$) of superior quality with half of the usual number of scans could be recorded on a 600Mhz spectrometer equipped with a cryoprobehead.

[f] $\tau_m = 100\mathrm{ms}$

The software package *sparky* (Goddard & Kneller, 2001) was used throughout the assignment process. It has a very valuable tool enabling the search for peaks at matching frequencies within one spectrum. At the ^{15}N and ^{1}H frequencies of a given amino acid, the CBCA(CO)NH and HNCA spectra show in the ^{13}C dimension the resonances for C^{α}_{i-1} and C^{β}_{i-1} or C^{α}_{i} and C^{α}_{i-1} respectively. Usually the 'own' C^{α}_{i} in the HNCA has a stronger intensity than the 'preceding' C^{α}_{i-1}. An automatic search for matching peaks for a given C^{α}_{i} will typically yield up to 30 hits, some of them being noise, others due to overlap. One of the hits will be the neighbor in the sequence, though, the C^{α}_{i} of the starting residue being the C^{α}_{i-1} of this one.

First of all, glycins were investigated. Out of the found possible neighbors, the one will be the real neighbor that shows only one peak in the CBCA(CO)NH spectrum. Any residue following a glycin in the sequence will lack the C^{β}_{i-1} resonance in the CBCA(CO)NH, as there is no C^{β} in glycins.

In a similar way, but with more chance for ambiguities, all other identified amino acids were investigated to find their neighbors and eventually short sets of residues with correlated shifts. These sets could then be tracked in the sequence of the protein to to arrive finally at the complete sequence specific (^{1}H, ^{15}N, ^{13}C) resonance assignment.

Figure 6.2 shows some strips from the CBCA(CO)NH and HNCA spectra. The assignment-walk through residues K75 to S79 is indicated by lines connecting the inter and intra C^{α} resonances. It can be seen, that the CBCA(CO)NH spectrum serves as a back up to verify the correct assignment of the C^{α}_{i-1} resonance and additionally provides the C^{β}_{i-1} frequencies for each amino acid.

Ambiguities which frequently arouse due to signal overlap could in some cases be resolved by looking at the strips from the ^{15}N-NOESY-HSQC spectrum. This spectrum was of superior quality as it was recorded using a cryo-probehead. In this kind of a probehead, the preamplifier and radio frequency coils are cooled to liquid helium temperatures, which increases the signal to noise ratio by a factor of 3-4[1].

CAP-N exhibits a mostly α helical fold in which the through space distance of two neighboring backbone amide protons is about 2.8 Å (Wüthrich, 1986). Therefore, for a great number of residues, H^{N}_{i}-H^{N}_{i-1} cross peaks could be found in the amide region of the 3D-NOESY. Some representative strips from the ^{15}N-NOESY-HSQC spectrum are shown in Figure 6.3. The discovered sequential cross-peaks aided not only the assignment, but represent valuable distance constraints used later in the structure calculations.

[1] I would like to thank Dr. Helena Kovacs of Bruker BioSpin, Fällanden, CH for recording this spectrum on the Cryoprobe^{TM} system.

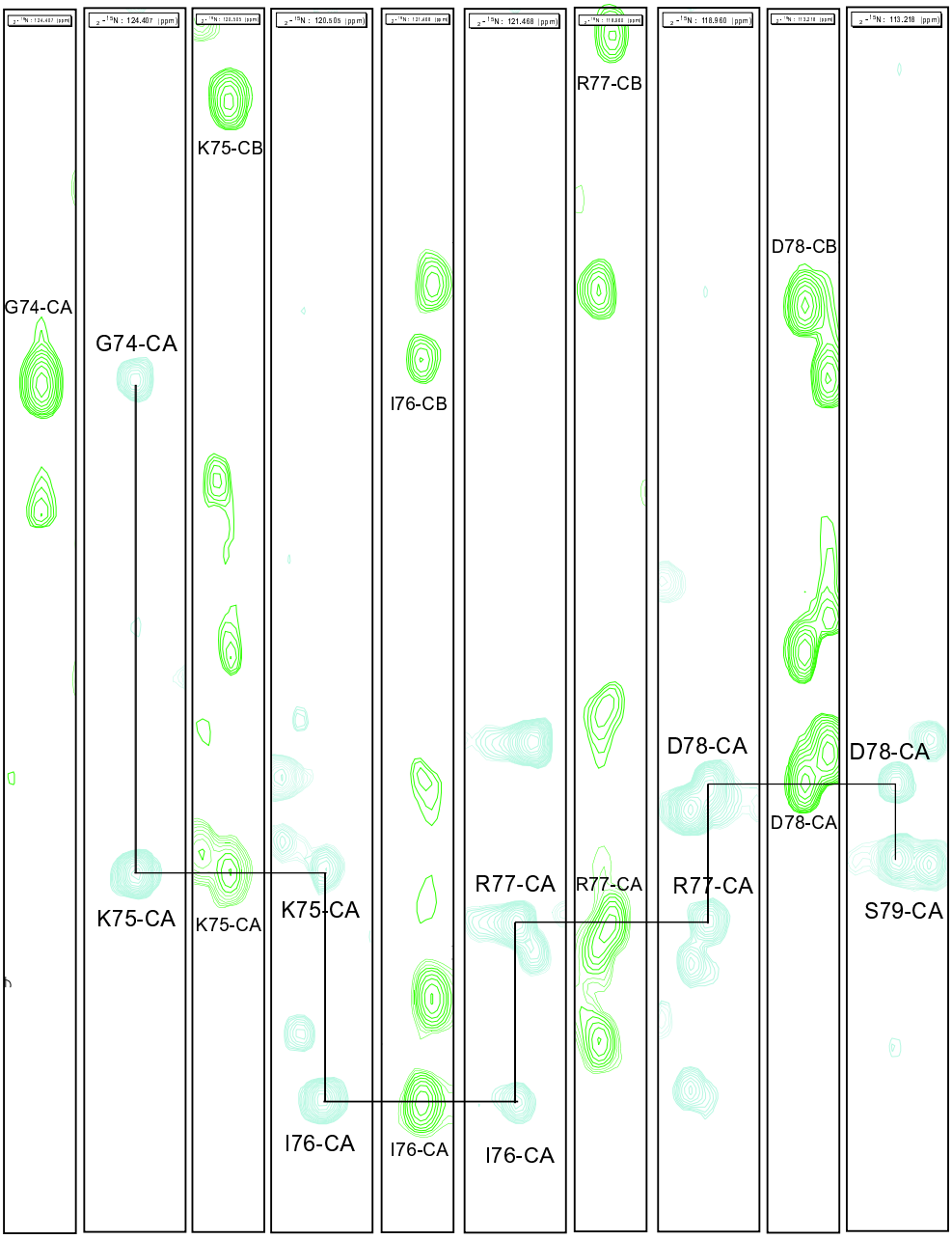

Figure 6.2: Strips from the CBCA(CO)NH (narrow strips, green) and HNCA (wide strips, turquoise) spectra. The sequential assignment is indicated.

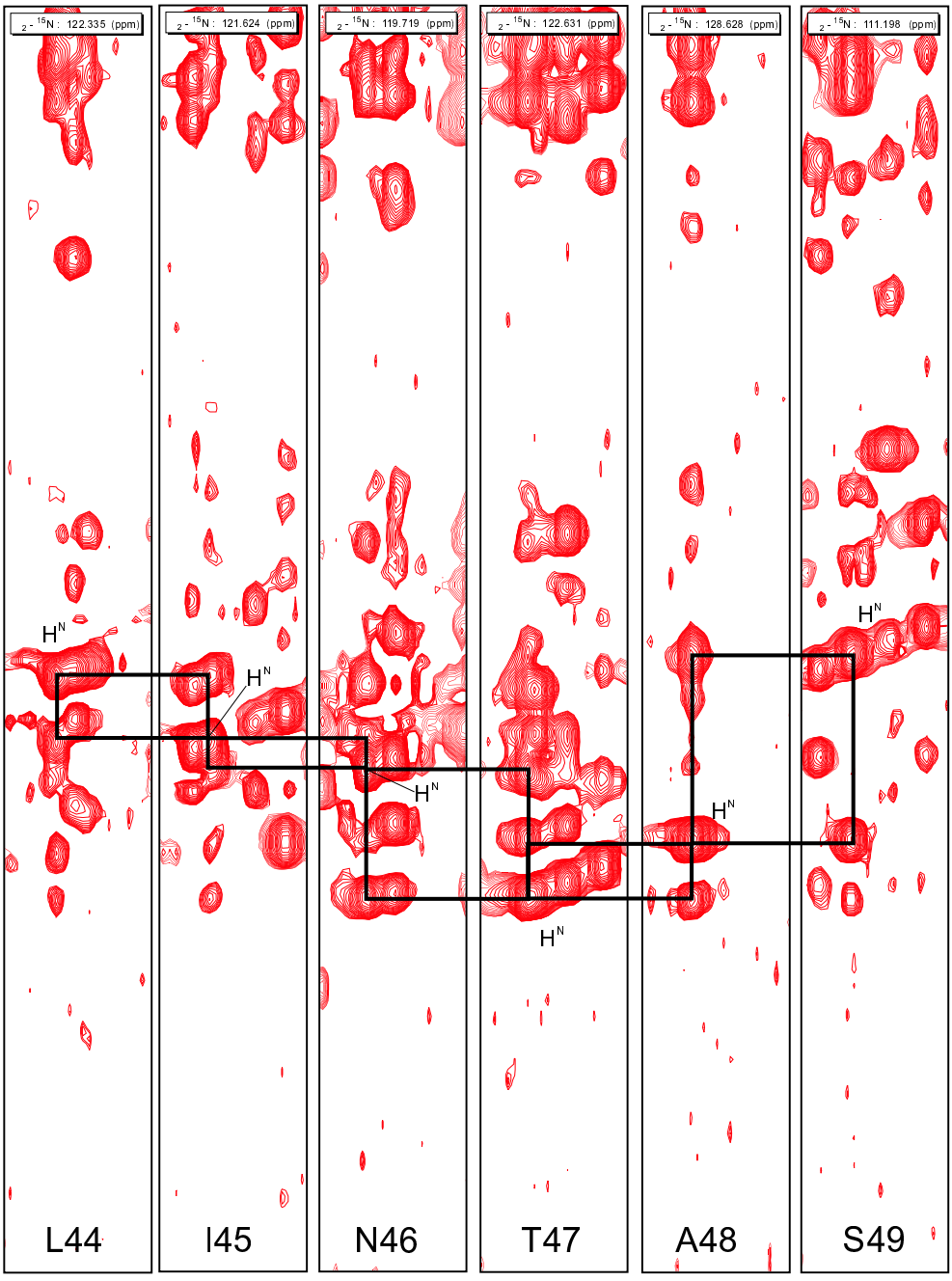

Figure 6.3: Strips from the ^{15}N-NOESY-HSQC spectrum. The connectivities between the backbone H^N diagonal- and cross-peaks of residues L44 through S49, which are part of the second α-helix, are indicated by black boxes. Diagonal peaks are labeled H^N in each strip.

After finishing the sequence specific assignment of the backbone amide groups and the C^{α} and C^{β} resonances, a HNCO spectrum was used to find the C' frequencies. As for the other ^{1}H-^{15}N-HSQC based triple resonance spectra, no signals of residues preceding prolines will be detected. Thus it is not possible to assign the C^{α}_{P-1}, C^{β}_{P-1} or C'_{P-1} residues of proline preceding residues while the C^{α}_{P}, C^{β}_{P} or C'_{P} of prolines itself can be assigned. ^{15}N-NOESY-HSQC and ^{13}C-NOESY-HSQC spectra were then used to assign as many side chain resonances as possible. This would later be the basis for finding NOE distance constraints for the three dimensional structure calculations.

Extent of assignment and data deposition

Figure 6.4 shows the ^{1}H-^{15}N-HSQC spectrum of CAP from *Dictyostelium discoideum*. Resonances of all backbone amide groups were assigned with the exception of the N-terminal residues S1 through K3, which could not be identified in the NMR spectra. Furthermore in the backbone 95% of H^{α} and 93% of C^{α} and C' were assigned as well as about half of the side chain atoms, including 95% of C^{β} and 81% of H^{β}. This assignment is sufficient to determine the structure of the protein and to analyze its dynamics. A table of the ^{1}H, ^{15}N, ^{13}C chemical shift assignment of CAP has been deposited in the BioMagResBank database (http://www.bmrb.wisc.edu) under the accession number 5393.

6.5 Three-dimensional structure determination

At the same time that the NMR structural determination got under way, the crystall structure was solved independently in our group (Ksiazek, 2002, see Figure 6.5). A substantial number of additional long range NOEs could be assigned in the NMR spectra on the basis of the interproton distances derived from the coordinates of this X-ray structure. The presence of these NOEs proofs, that the solution and crystall structures are very similar. Original calculations based on NMR data included roughly 500 mostly sequential and medium range NOEs and 38 ϕ-angles in the backbone derived from an HNHA experiment (see Table 6.2). These calculations revealed the totaly α helical secondary structure of the protein that was later confirmed by the X-ray structure. The tertiary fold could not be established with NMR data alone.

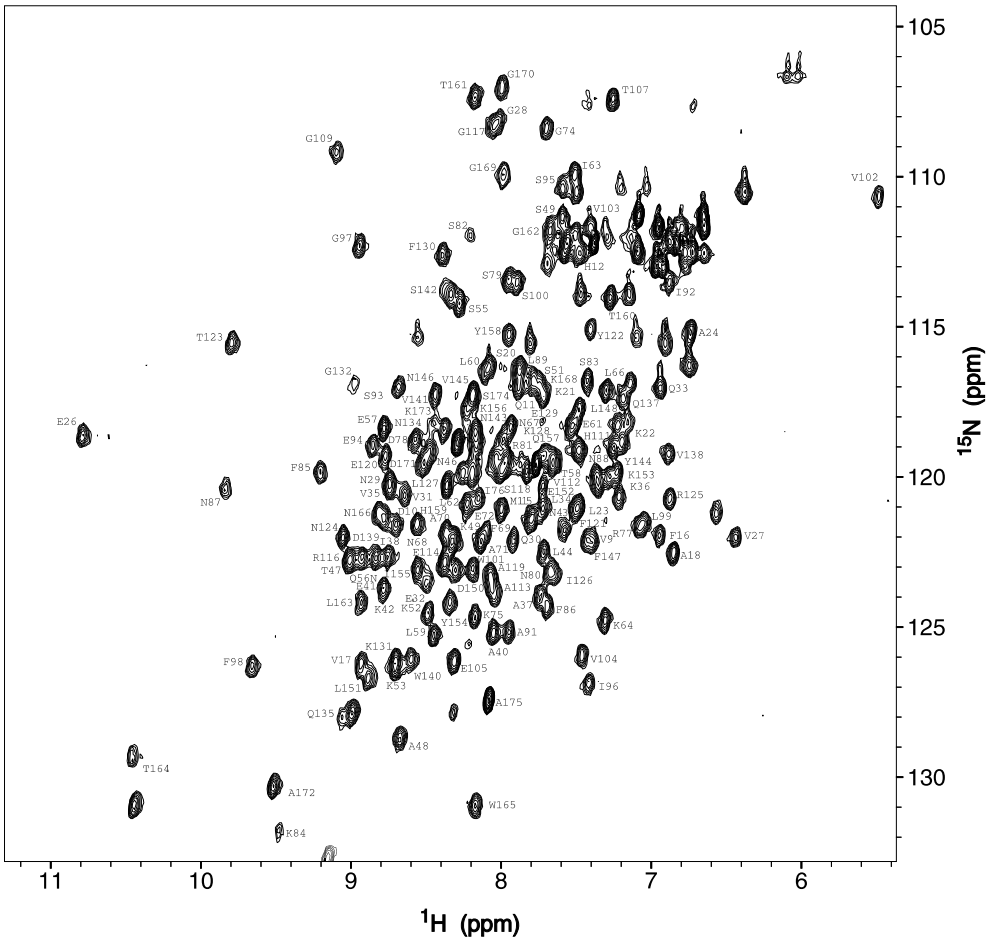

Figure 6.4: The ^{1}H-^{15}N-HSQC spectrum of CAP-N from *Dictyostelium discoideum* at 300K and pH 7.3. The residue specific resonance assignment is indicated by the one letter amino acid code next to the corresponding signal; several are omitted for clarity.

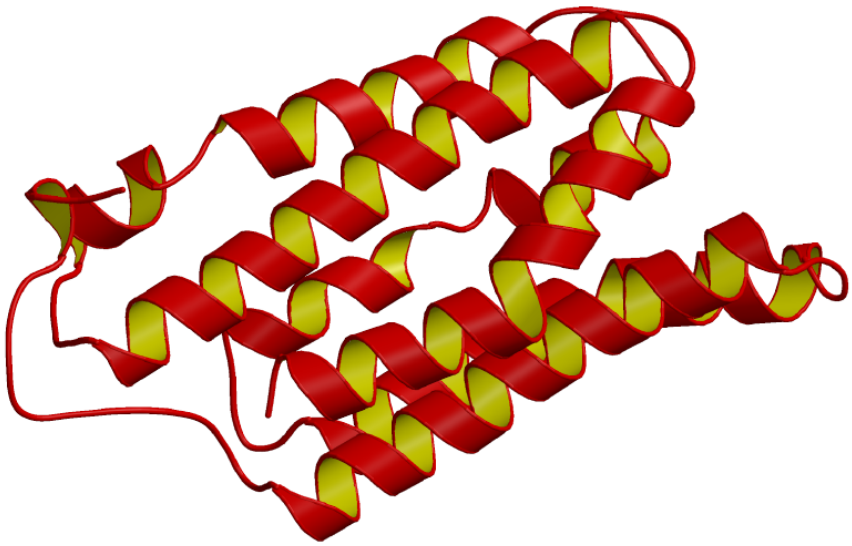

Figure 6.5: Ribbon drawing of the three-dimensional structure of CAP-N as revealed by X-ray crystallography (Ksiazek, 2002).

6.6 ^{15}N-Relaxation

Protein function is closely related not only to the three-dimensional structure but also to intramolecular motions. As NMR relaxation studies can provide detailed, residue per residue, information on internal dynamics of proteins, they are routinely applied to ^{15}N-labeled samples (for a review on protein NMR relaxation see Fischer et al. (1998)).

Typically, three relaxation parameters are measured for backbone amide groups of proteins: the longitudinal relaxation time T_1, the transverse relaxation time T_2 and the heteronuclear Overhauser enhancement {^{1}H}-^{15}N-NOE.

The N-terminal domain of CAP under investigation here is a very rigid protein. Apart from the C-terminal 10 residues, the relaxation data showed no increased flexibility for any part of the protein. The C-terminus, which is not structured, is highly flexible as seen from the {^{1}H}-^{15}N-NOE and T_2 data. This is especially true for residues S174 to T176, which also show no long range NOESY contacts. The T_1 and T_2 values determined here are odd for a protein with 176 residues and may suggest dimerization (see section 6.7 below).

T_1 Relaxation

^{15}N T_1 measurements were recorded with T_1 relaxation delays of 12.4, 384.4, 756.4, 1118.4 and 1500.4 ms. Six spectra were recorded in an interleaved manner, the first and last one having the same delay of 12.4 ms. The peak intensities were fit to a decaying exponential using the python relax.py extension (rh) in sparky (Goddard & Kneller, 2001).

Figure 6.6 shows the T_1 values for CAP-N. Estimated errors result mainly from signal overlap. The determined T_1 times are rather uniform, showing no region of significantly reduced or enhanced relaxation times. The average T_1 time for the whole protein is 920.5 ms.

T_2 Relaxation

The ^{15}NT_2 times were determined analog to the T_1 measurements. T_2 relaxation delays were set to 20.8, 41.6, 83.2, 125.0 and 166.0 ms. Figure 6.7 shows the T_2 values for CAP-N. The flexible C-terminus of the protein is clearly visible. Apart from it, no enhanced T_2 times were determined. The average T_2 time for the whole protein is 54.0 ms and 50.9 if the C-terminal 10 residues, which are not structured, are neglected.

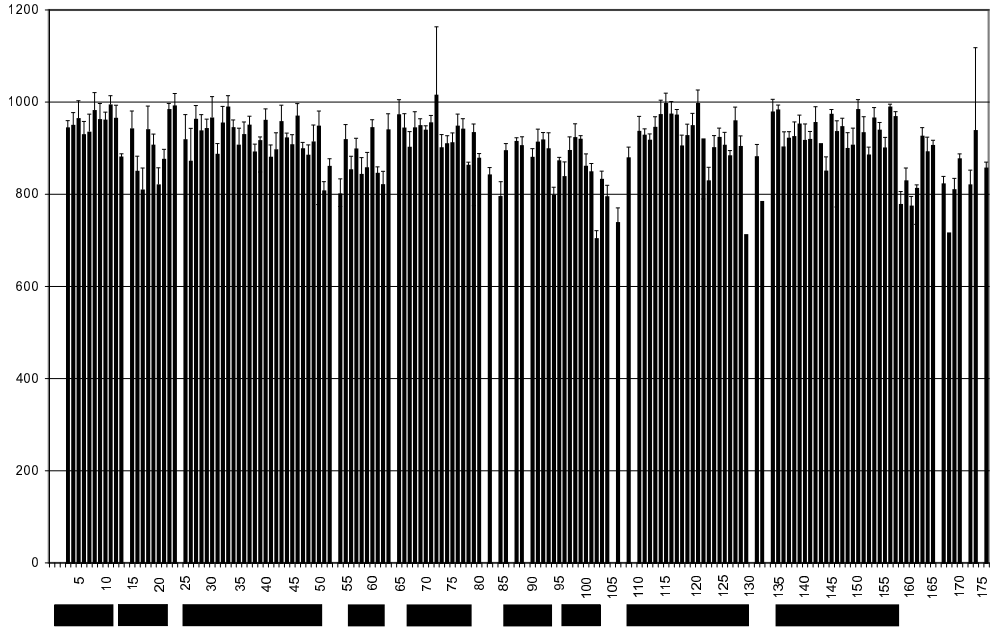

Figure 6.6: ^{1}H T_1 relaxation times ploted for the assigned residues of CAP-N. Blanks correspond to prolines in the sequence. The positions of α-helices as revealed by the x-ray structure (Ksiazek, 2002) are indicated as black boxes.

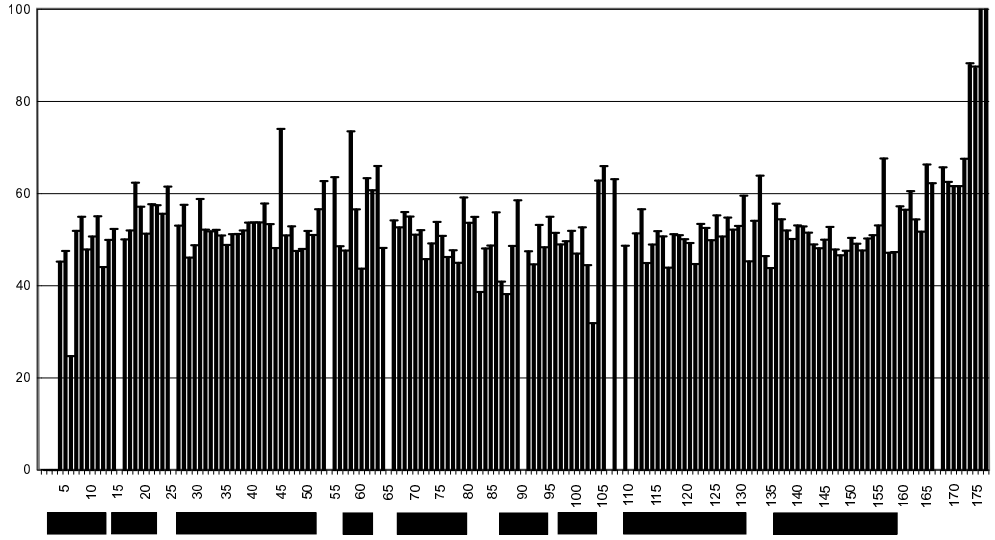

Figure 6.7: ^{1}H T_2 relaxation times of CAP-N. Values for A175 and T176 (off scale) were 150.2$\pm$15.3 and 334.4$\pm$16.8, respectively; α-helices are indicated.

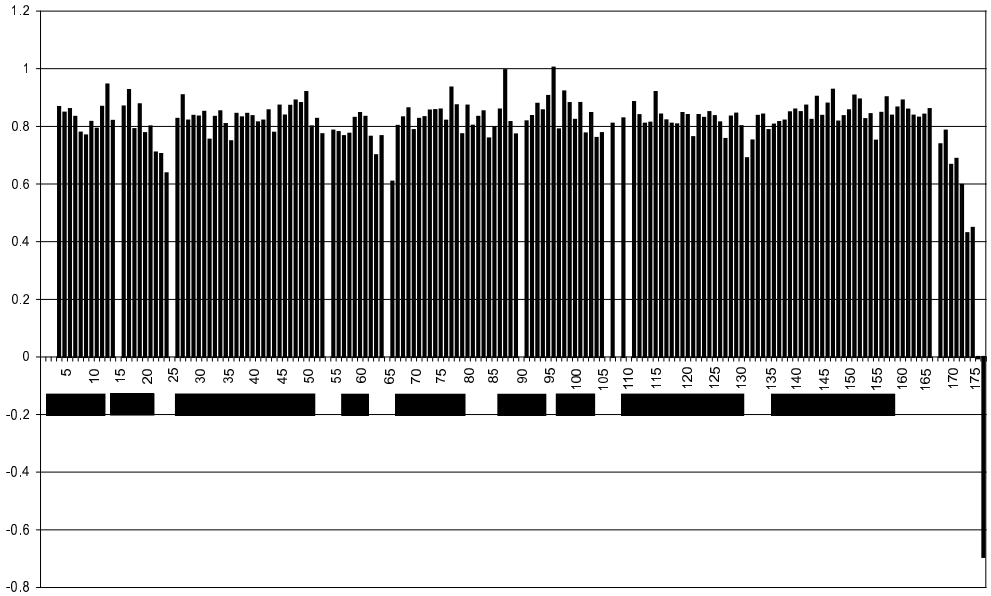

Figure 6.8: Heteronuclear $\{^1H\}$-^{15}N-NOE values of CAP-N; α-helices are indicated.

$\{^1H\}$-^{15}N-NOE

The heteronuclear Overhauser enhancement values were determined using a pair of spectra with and without proton saturation during the recycle delay. Saturation was achieved by a train of 120° pulses separated by 10 ms for a time of 3 s. Additionally a 4.5 s relaxation delay was used between scans. The two spectra were recorded in an interleaved manner to avoid time-dependent artifacts. Other spectral parameters were identical to the HSQC experiments (Table 6.2).

The calculated NOE values are shown in Figure 6.8. The average value for the whole protein is 0.80 and slightly higher neglecting the C-terminal 10 residues (0.83). Some regions show reduced NOE values. These residues (K22-A24, I63 and L66, K131) are part of the loops that connect the helices. No NOE values are below 6.0 apart from the C-terminus, with A175 close to zero and T176 clearly negative. CAP-N thus is very rigid on timescales $<$ 10 ns.

6.7 Is CAP a Dimer?

Much has been speculated on whether members of the CAP family dimerize. For CAP from *Saccharomyces cervisiae* Hubberstey & Mottillio (2002) propose a N-terminal region close to

the adenylyl cyclase binding site to be important for dimerization. Interestingly enough, this is exactly the region proved to be unstructured in *Dictyostelium* CAP and this fragment was removed for this investigations (see section 6.2). This could be the reason why *Dictyostelium* CAP, as CAPs from other higher eukaryotes, apparently does not bind adenylyl cyclase (Hubberstey & Mottillio, 2002). Attempts to crystallize CAP-N in our group lead to two different crystall structures, one being a monomer (see Figure 6.5) the other a dimer (see Figure 6.9). The residues involved in dimerization were identified to be K64, N67, A71, D78, R81, E94, M115, F121, Y122 and R125 (Ksiazek, 2002). The NMR relaxation data reported in section 6.6 also suggests, that CAP-N is a dimer in solution. The determined average T_1 and T_2 values (920.5 ms and 54.0ms) are far too long and too short, respectively, for CAP-N to be a monomer. The mass spectra mentioned in section 6.2 also show a clear peak at twice the calculated mass of CAP-N. Thus CAP-N is a mixture of monomers and dimers at concentrations around 1 mM, pH 7.3 and 300K in solution as used for NMR investigations.

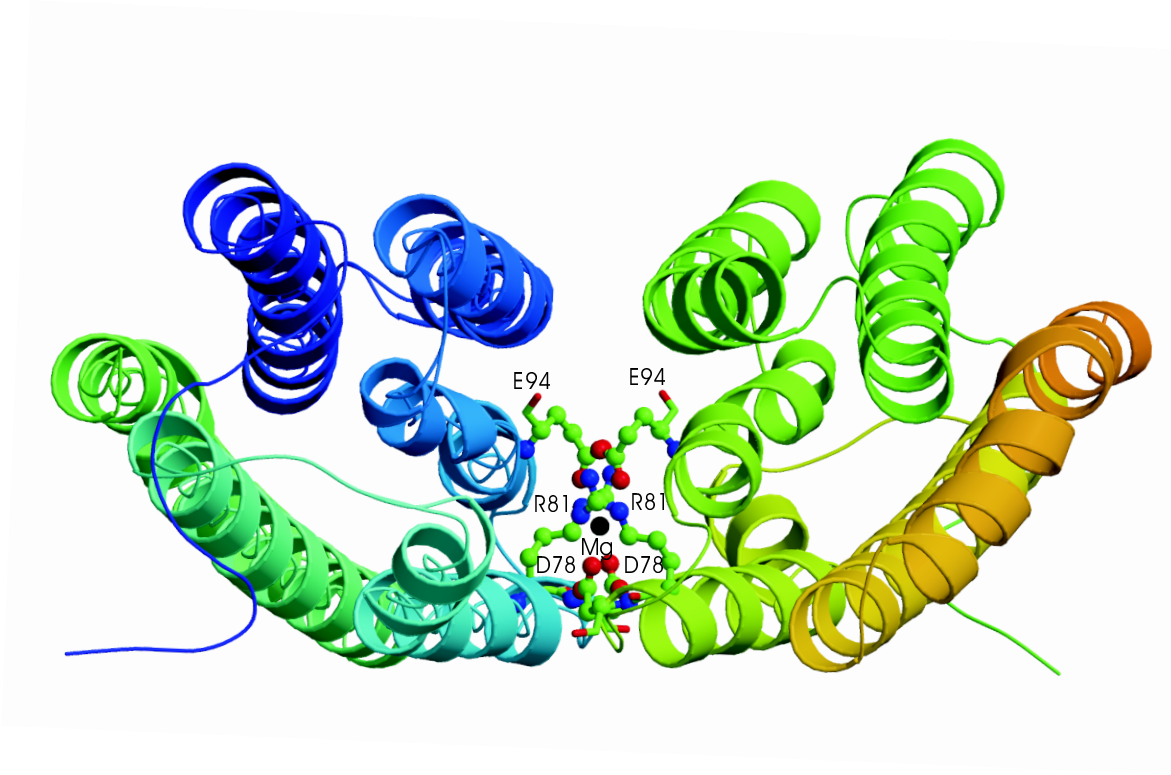

Figure 6.9: Ribbon drawing of the three-dimensional structure of the CAP-N dimer as revealed by X-ray crystallography (Ksiazek, 2002). In the center of the dimerization site, a magnesium atom is present.

Chapter 7

NMR Characterization of the Green Fluorescent Protein

7.1 Biological context

Green fluorescent proteins (GFPs) provide powerful tools for monitoring gene expression, protein movement and protein interaction (Tsien, 1998). Due to its inherent fluorescence GFP is a widely used fusion tag for proteins in fluorescence microscopy studies. Up to now the family of GFP-like proteins comprises 27 cloned and spectroscopically characterized proteins (Labas et al., 2002) of which 18 high-resolution crystal structures are available (www.rcsb.org; Berman et al. (2000)).The overall structure of GFP consists of an 11-stranded β-barrel with a centrally located helix that carries the chromophore (Ormoe et al., 1996; Yang et al., 1996). The X-ray diffraction studies and a variety of physicochemical methods highlight an apparent exceptional stability of the GFP fold in which the chromophore lies rigidly inside the conformationally inflexible GFP molecule (Striker et al., 1999). The stability of GFP against temperature, denaturants and proteases is very high (Tsien, 1998; Ward, 1981). Recently ^{19}F NMR studies of the cyan variant of GFP conducted in our laboratory indicated conformational flexibility in or near the chromophore moiety with residue His148 being most likely involved in this process (Seifert et al., 2002). NMR has been suggested as a tool for elucidating the dynamics of chromophore formation, water accessibility of the chromophore and conformational flexibility in GFP (Prendergast, 1999; Haupts et al., 1998). In contrast to many spectroscopic techniques, NMR spectroscopy provides a large frequency range for studying dynamical processes from picosecond to second timescales and even longer at atomic resolutions. For example, motional and thermodynam-

ical information for backbone amides in proteins can be obtained from measurements of ^{15}N relaxation rates in ^{15}N labeled proteins (Spyracopoulos & Sykes, 2001; Kay, 1998). Hydrogen-deuterium exchange experiments allow for a characterization of conformational fluctuations in secondary structure elements (Dempsey, 2001). The basis of these relaxation measurements which were subsequently performed in our laboratory (Seifert, 2002) is the sequence-specific resonance assignment of the protein spectra.

7.2 NMR Spectroscopy

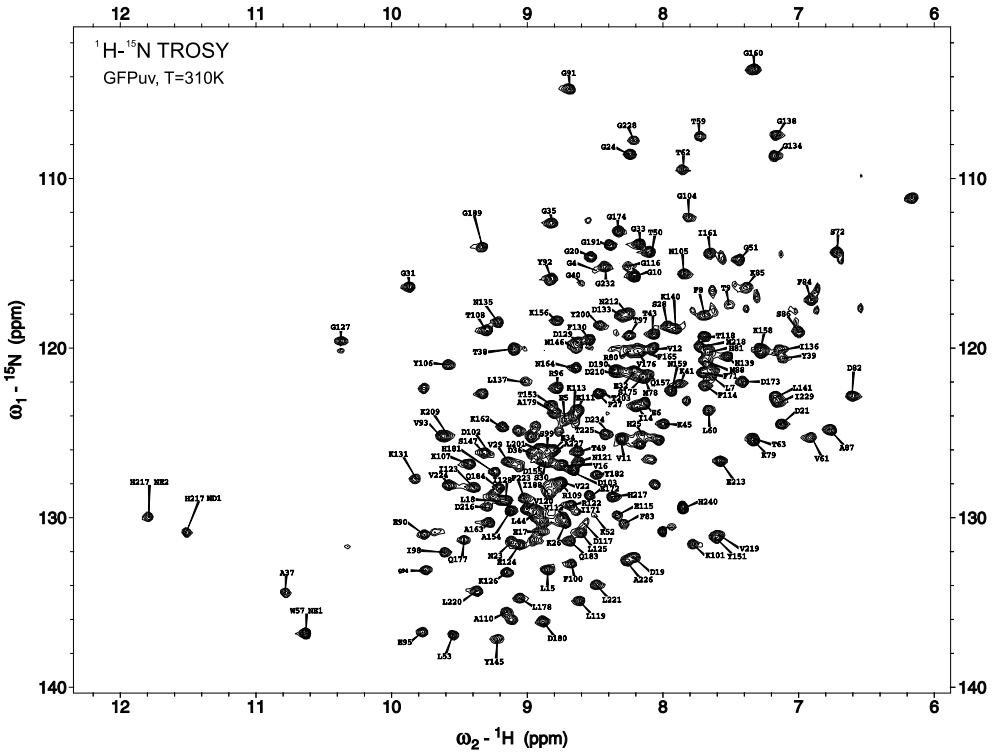

Figure 7.1: 2D ^{1}H-^{15}N TROSY spectrum of GFPuv at 310K. The labels indicate the sequence-specific resonance assignment.

Materials

^{15}N and ^{15}N/^{13}C labeled samples of GFPuv (Q80R, F99S, M153T, V163A, also known as 'cycle 3' (Crameri et al., 1996)) were produced as described by Georgescu (2000). The samples were

dissolved in PBS (115 mM NaCl, 8 mM KH_2PO_4, 16 mM Na_2HPO_4) buffer at pH 7.3. The following samples were available in concentrations ranging from 0.8 to 1.2 mM: Uniformly ^{15}N labeled GFPuv, as well as selectively labeled ^{15}N-Ala, ^{15}N-Phe, ^{15}N-Gly, ^{15}N-Ile, ^{15}N-Lys, ^{15}N-Leu, ^{15}N-Val, ^{15}N-Tyr and ^{15}N-Met and reverse labeled ^{15}N-Asn, ^{15}N-Thr, ^{15}N-His and ^{15}N-Asp samples. Additionally a ^{15}N-^{13}C double labeled sample and ^{15}N/99%^{2}H and ^{15}N/^{13}C/70%^{2}H deuterated samples were prepared. All samples contained 10% D_2O.

Backbone H^N, N, C^α, and C^β assignment of the GFPuv mutant

All heteronuclear ^{1}H-^{15}N NMR experiments were performed on Bruker DMX 750, DRX 600 and DRX 500 spectrometers at a temperature of 310K. The spectrometers were equipped with 5 mm ^{1}H-^{13}C-^{15}N triple resonance TXI probeheads including triple-axis gradients (DMX 750, DRX600) or z-axis gradients (DRX500). The spectra were recorded with a sweepwidth of 17.5 ppm for ^{1}H and 42 ppm for ^{15}N. The backbone resonances were assigned using HNCO, HN(CO)CA and a pair of HNCA and CBCA(CO)NH triple resonance spectra (Cavanagh et al., 1996). In addition ^{15}N-HSQC spectra (Mori et al., 1995) of the selectively and uniformly ^{15}N labeled samples, 2D-TROSY, 2D-NOESY ($\tau_m = 160$ ms), and 3D-NOESY-HSQC ($\tau_m = 160$ ms) spectra (Cavanagh et al., 1996; Salzmann et al., 1998) were recorded at various field strength. The assignment was accomplished using the program *Sparky* (Goddard & Kneller, 2001) and the in house software package *ccnmr* (Cieslar et al., 1988, 1990).

Extent of assignment and data deposition

In summary more than 80% (191 out of 229) of the backbone amide groups of GFPuv were assigned. Also assignment of the C^α, C^β and C' atoms of these residues was obtained. The assignment of the central α-helix and of β-sheets 7, 8, and 10 could not be completed. These β-sheets are assumed to form the interface in dimerization of the protein. It also turned out that the backbone assignment was complicated by the necessity of a high pH and the tendency of GFP to aggregate. The chemical shifts have been deposited in the BioMagResBank database under the accession number 5514 (http://www.bmrb.wisc.edu).

Chapter 8

Summary

This thesis is a collection of several NMR (nuclear magnetic resonance spectroscopy) projects conducted at the Department of Structural Research of the Max Planck Institute of Biochemistry in the recent years. Each chapter represents a project that yielded a paper as given on page *i*. Some of these still have to be published.

The wide range of applications of NMR in biochemical research is reflected in the diversity of the projects. One of these applications, the use of NMR as a screening tool in structural proteomics is reviewed. Using several examples from our group, it is shown that NMR can serve as a powerful tool to identify protein samples that are suitable for structure elucidation by both NMR spectroscopy and X-ray crystallography.

Two projects concentrated on the detection of ligands for proteins that play a major role in diseases including cancer. In the first case it could be shown that chalcone derivatives (compounds derived from 1,3- diphenyl-2-propen-1-one) are inhibitors of MDM2 (human murine double minute clone 2 protein). In some cancers overexpression of MDM2 inactivates an otherwise intact p53, disabling the genome integrity checkpoint and allowing cell cycle progression of defective cells. MDM2 binds to the transactivation domain of p53. The chalcones described here bind to a subsite of the p53 binding cleft of human MDM2 and thus disrupt the interaction of MDM2 with p53. Binding studies were performed by titrating MDM2 with increasing amounts of the chalcones and recording heteronuclear NMR spectra of each step. Changes of the chemical shift of the backbone amide groups indicate the binding site of the inhibitors as well as the binding strength.

In a methodologically similar project, inhibitors of the IGF-I (insulin-like growth factor-I) and IGF-binding protein-5 (IGFBP-5) interaction were targeted. In this case an *in silco* screen-

ing with the software FlexX had suggested a number of possible inhibitors which were subsequently investigated by NMR. One compound, an FMOC derivative (a derivative from 9-Fluorenylmethoxycarbonyl) was found to bind to IGFBP-5 weakly. This compound was then optimized to increase its affinity. The final designed ligand had a considerably lower K_D value but was still not able to disrupt the IGF/IGFBP-5 complex. Still it could be the starting point for the design of more potent inhibitors and therapeutic agents for diseases that are associated with abnormal IGF-I regulation like some cancers, but also stroke and other neurodegenerative diseases.

NMR-based structural studies and NMR-based ligand binding studies require selectively and/or uniformly ^{15}N-labeled proteins. The baculovirus expression system based on insect cells has several advantages over commonly used bacterial systems. To achive labeling in the baculovirus expression system, two insect media for the production of selectively labeled protein in insect cells were developed in our group. Again heteronuclear experiments were used to validate the quality of the media and to investigate the metabolic pathways in Sf9 insect cells. As a result, now two valuable media are about to be published (previous compositions were kept confidential) and a preliminary model of the metabolism of Sf9 cells could be presented.

In two more structurally oriented projects, NMR characterizations of the globular proteins GFP (green fluorescent protein) and CAP (cyclase associated protein) were performed. While in the first case the focus was on dynamical properties, the ultimate goal of the investigations on CAP was the discovery of its three dimensional structure. A nearly complete resonance assignment of the protein was achieved and structure calculations were under way, when the structure was solved by X-ray crystallography. The secondary structure elements of the X-ray structure are in agreement with those predicted by NMR spectroscopy.

Chapter 9

Zusammenfassung

Diese Dissertation stellt eine Sammlung verschiedenster Projekte dar die in den letzten Jahren in der Abteilung für Strukturforschung am Max-Planck-Institut für Biochemie mittels der Kernresonanzspektroskopie durchgeführt wurden. Jedes Kapitel repräsentiert ein Projekt, das zu einem der auf Seite *i* angegebenen Artikel geführt hat. Einige sind allerdings noch nicht veröffentlicht.

Die große Bandbreite der Anwendungen der Kernresonanzspektroskopie (nuclear magnetic resonance spectroscopy, NMR) in der biochemischen Forschung spiegelt sich auch in den sehr unterschiedlichen Projekten wieder. Eine dieser Anwendungen wurde in einem Übersichtsartikel, der am Anfang dieser Arbeit wiedergegeben ist genauer betrachtet. An Hand einiger Beispiele aus unserer Arbeitsgruppe wurde gezeigt, wie die NMR-Spektroskopie dazu beitragen kann solche Proteine auszuwählen, die für eine dreidimensionale Strukturbestimmung mittels NMR-Spektroskopie oder Röntgenstrukturanalyse geeignet sind.

Zwei der Projekte befassten sich mit der Identifizierung von Bindungspartnern von Proteinen die in schweren Krankheiten, u.a. auch Krebs, eine Rolle spielen. In einem Fall konnte gezeigt werden, dass Chalcone-Derivate (Derivate der chemischen Verbindung 1,3-diphenyl-2-propen-1-one) MDM2 (human murine double minute clone 2 protein) inhibieren. In einigen Krebsarten inaktiviert MDM2 das Protein p53 und ermöglicht so eine Fortlaufen des Zellzykluses von kranken Zellen. MDM2 bindet dabei an die Transaktivierungsdomäne von p53. Die hier beschriebenen Chalcone binden an einen Teil der p53 Bindungstasche des humanen MDM2 und verhindern so eine Bindung zwischen MDM2 und p53. Für die Bindungsstudien wurde MDM2 mit wachsenden Mengen der Chalcone titriert und heteronukleare NMR Spektren von jedem Schritt aufgenommen. Veränderungen der chemischen Verschiebung der Amid-Gruppen

im Protein Rückgrat weisen sowohl auf die Bindungsstelle als auch auf die Dissoziationskonstante hin.

In einem methodisch ähnlichen Projekt wurden Inhibitoren der Bindung zwischen IGF-I (insulin-like growth factor-I) und dem IGF-bindenden Protein IGFBP-5 (IGF-binding protein-5) gesucht. In diesem Fall hatte eine computergestützte Suche mit dem Programm FlexX eine Reihe vom möglichen Bindungspartnern für IGFBP-5 ergeben, die dann ebenfalls mit der Kernresonanzspektroskopie weiter untersucht wurden. Eines der Moleküle, ein FMOC-Derivat (Derivat der chemischen Verbindung 9-Fluorenylmethoxycarbonyl) zeigte eine schwache Bindung an IGFBP-5. Es wurde daraufhin verändert um seine Affinität zu steigern. Am Ende wurde so ein Ligand gefunden, der eine deutlich niedrigere Dissoziationskonstante aufwies, allerdings nicht in der Lage war den Komplex aus IGF und IGFBP-5 aufzubrechen. Dieses Molekül könnte aber als Ausgangspunkt für die Suche nach Inhibitoren dienen um Medikamente für einige Krebsarten die mit einer fehlerhaften IGF-I Regulation einhergehen, aber auch für Schlaganfälle und andere Hirnverletzungen, zu finden.

Viele strukturbezogene Untersuchungen mit Hilfe der Kernresonanzspektroskopie erfordern selektiv und/oder komplett mit ^{15}N markierte Proteinproben. In vielen Fällen hat es sich als vorteilhaft erwiesen Proteine in einem Insekten-Zellkultursystem, dem sogenannten Baculovirus-System, zu exprimieren, da sowohl die Ausbeute als auch die Faltung der Proteine oft besser sind als im üblicherweise verwendeten *E.coli*-System. Um auch im Baculovirus Expressionssystem markieren zu können, wurden in unserer Gruppe zwei Medien entwickelt, mit denen in Insektenzellen ^{15}N Markierung möglich ist. Wiederum wurden heteronucleare NMR Experimente verwendet um die Qualität der Medien zu überprüfen und den Metabolismus von Sf9 Insektenzellen zu untersuchen. Diese zwei Medien sind nun der Öffentlichkeit zugänglich (bereits existierende Rezepte wurden bisher stets geheim gehalten) und ein vorläufiges Modell des Metabolismus von Sf9-Zellen konnte vorgestellt werden.

In zwei mehr strukturorientierten Projekten wurden die globulären Proteine GFP (green fluorescent protein) und CAP (cyclase associated protein) mit Hilfe der NMR Spektroskopie charakterisiert. Während bei ersterem die dynamischen Eigenschaften im Vordergrund der Untersuchungen standen, sollte bei CAP die dreidimensionale Struktur des Proteins bestimmt werden. Dabei wurde ein Großteil der Signale in den Spektren den einzelnen Atomen des Proteins zugewiesen und erste Strukturrechnungen begonnen. Mittlerweile wurde die Struktur allerdings mit Hilfe dere Röntgenstrukturanalyse aufgeklärt. Die Sekundärstrukturelemente der Röntgenstruktur decken sich weitgehend mit den mittels NMR-Spektroskopie getroffenen

Vorhersagen.

Appendix A

BioMagResBank entry for CAP-N

Sequence-specific (^{1}H,^{15}N ,^{13}C) resonance assignment of the N-terminal domain of the cyclase-associated protein (CAP) from Dictyostelium discoideum as deposited at the BioMagResBank database (http://www.bmrb.wisc.edu) under the accession number 5393.

```
data_5393

######################
#  Entry information  #
######################


save_entry_information
   _Saveframe_category        entry_information

   _Entry_title
;
Sequence-specific (1H, 15N, 13C) resonance assignment
of the N-terminal domain of the Cyclase-associated
Protein (CAP) from Dictyostelium discoideum
;

   loop_
      _Author_ordinal
      _Author_family_name
      _Author_given_name
      _Author_middle_initials
      _Author_family_title

       1   Rehm        Till        .  .
       2   Mavoungou   Chrystelle  .  .
       3   Israel      Lars        .  .
       4   Schleicher  Michael     .  .
       5   Holak       Tad         A  .

   stop_

   _BMRB_accession_number    5393
   _BMRB_flat_file_name      bmr5393.str
   _Entry_type               new
   _Submission_date          2002-06-12
   _Accession_date           2002-06-13
   _Entry_origination        author
   _NMR_STAR_version         2.1.1
   _Experimental_method      NMR

   loop_
      _Saveframe_category_type
      _Saveframe_category_type_count

       assigned_chemical_shifts  1

   stop_

   loop_
      _Data_type
      _Data_type_count

      '1H chemical shifts'   714
      '15N chemical shifts'  173
      '13C chemical shifts'  542

   stop_

   _Release_immediately      no
   _Release_on_publication   yes

save_


save_contact_persons
   _Saveframe_category    contact_persons


   loop_
      _Family_name
      _Given_name
      _Middle_initials
      _Family_title
      _Department_and_Institution
      _Mailing_address
      _Phone_number
      _Email_address
      _FAX_number

       Rehm
```

```
       Till
       .
       .

;
Dept. of Structural Research
Max-Planck-Institute of Biochemistry
;

;
am Klopferspitz 18A
82152 Martinsried
Germany
;
      "+49 89 8578 2672"
       rehm@biochem.mpg.de
      "+49 89 8578 3777"

       Holak
       Tad
       A
       .

;
Dept. of Structural Research
Max-Planck-Institute of Biochemistry
;

;
am Klopferspitz 18A
82152 Martinsried
Germany
;
      "+49 89 8578 2673"
       holak@biochem.mpg.de
      "+49 89 8578 3777"


   stop_

save_


##############################
#  Citation for this entry  #
##############################


save_entry_citation
   _Saveframe_category     entry_citation

   _Citation_title
;
Letter to the Editor: Sequence-specific (1H, 15N, 13C)
resonance assignment of the N-terminal domain of the
Cyclase-associated Protein (CAP) from Dictyostelium
 discoideum
;
   _Citation_status        submitted
   _Citation_type          journal
   _MEDLINE_UI_code        .

   loop_
      _Author_ordinal
      _Author_family_name
      _Author_given_name
      _Author_middle_initials
      _Author_family_title

       1   Rehm        Till        .  .
       2   Mavoungou   Chrystelle  .  .
       3   Israel      Lars        .  .
       4   Schleicher  Michael     .  .
       5   Holak       Tad         A  .

   stop_

   _Journal_abbreviation  "J. Biomol. NMR"
   _Journal_volume         ?
   _Journal_issue          ?
   _Page_first             ?
   _Page_last              ?
   _Year                   ?

   loop_
      _Keyword

      "adenylyl cyclase associated protein"

   stop_

save_


##################################
#  Molecular system description  #
##################################


save_system_CAP
   _Saveframe_category         molecular_system

   _Mol_system_name
"N-terminal domain of the adenylyl cyclase associated Protein"
   _Abbreviation_common        CAP
   _Enzyme_commission_number   .

   loop_
      _Mol_system_component_name
      _Mol_label

      "CAP N-terminus" CAP

   stop_

   _System_physical_state      native
   _System_oligomer_state      monomer
   _System_paramagnetic        no
   _System_thiol_state        'not present'

save_


    ########################
    #  Monomeric polymers  #
    ########################


save_CAP
   _Saveframe_category     monomeric_polymer

   _Mol_type               polymer
   _Mol_polymer_class      protein
   _Name_common           "Cyclase associated Protein"
```

```
   _Name_variant           .
   _Abbreviation_common    CAP
   _Mol_thiol_state       'not present'
   _Details
;
The N-terminal domain of CAP described here is shorter by 39 residues
than previously reported.
;

    ##############################
    #  Polymer residue sequence  #
    ##############################


   _Residue_count          176
   _Mol_residue_sequence
;
SVKEFQNLVDQHITPFVALS
KKLAPEVGNQVEQLVKAIDA
EKALINTASQSKKPSQETLL
ELIKPLNNFAAEVGKIRDSN
RSSKFFNNLSAISESIGFLS
WVVVEPTPGPHVAEMRGSAE
FYTNRILKEFKGVNQDQVDW
VSNYVNFLKDLEKYIKQYHT
TGLTWNPKGGDAKSAT
;

   loop_
      _Residue_seq_code
      _Residue_label

   1   SER     2   VAL     3   LYS     4   GLU     5   PHE
   6   GLN     7   ASN     8   LEU     9   VAL    10   ASP
  11   GLN    12   HIS    13   ILE    14   THR    15   PRO
  16   PHE    17   VAL    18   ALA    19   LEU    20   SER
  21   LYS    22   LYS    23   LEU    24   ALA    25   PRO
  26   GLU    27   VAL    28   GLY    29   ASN    30   GLN
  31   VAL    32   GLU    33   GLN    34   LEU    35   VAL
  36   LYS    37   ALA    38   ILE    39   ASP    40   ALA
  41   GLU    42   LYS    43   ALA    44   LEU    45   ILE
  46   ASN    47   THR    48   ALA    49   SER    50   GLN
  51   SER    52   LYS    53   LYS    54   PRO    55   SER
  56   GLN    57   GLU    58   THR    59   LEU    60   LEU
  61   GLU    62   LEU    63   ILE    64   LYS    65   PRO
  66   LEU    67   ASN    68   ASN    69   PHE    70   ALA
  71   ALA    72   GLU    73   VAL    74   GLY    75   LYS
  76   ILE    77   ARG    78   ASP    79   SER    80   ASN
  81   ARG    82   SER    83   SER    84   LYS    85   PHE
  86   PHE    87   ASN    88   ASN    89   LEU    90   SER
  91   ALA    92   ILE    93   SER    94   GLU    95   SER
  96   ILE    97   GLY    98   PHE    99   LEU   100   SER
 101   TRP   102   VAL   103   VAL   104   VAL   105   GLU
 106   PRO   107   THR   108   PRO   109   GLY   110   PRO
 111   HIS   112   VAL   113   ALA   114   GLU   115   MET
 116   ARG   117   GLY   118   SER   119   ALA   120   GLU
 121   PHE   122   TYR   123   THR   124   ASN   125   ARG
 126   ILE   127   LEU   128   LYS   129   GLU   130   PHE
 131   LYS   132   GLY   133   VAL   134   ASN   135   GLN
 136   ASP   137   GLN   138   VAL   139   ASP   140   TRP
 141   VAL   142   SER   143   ASN   144   TYR   145   VAL
 146   ASN   147   PHE   148   LEU   149   LYS   150   ASP
 151   LEU   152   GLU   153   LYS   154   TYR   155   ILE
 156   LYS   157   GLN   158   TYR   159   HIS   160   THR
 161   THR   162   GLY   163   LEU   164   THR   165   TRP
 166   ASN   167   PRO   168   LYS   169   GLY   170   GLY
 171   ASP   172   ALA   173   LYS   174   SER   175   ALA
 176   THR

stop_

save_


    ####################
    #  Natural source  #
    ####################


save_natural_source
   _Saveframe_category   natural_source


   loop_
      _Mol_label
      _Organism_name_common
      _NCBI_taxonomy_ID
      _Superkingdom
      _Kingdom
      _Genus
      _Species

      $CAP
      "Dictyostelium discoideum"
       44689
       Eukaryota
       .
       Dictyostelium
       discoideum


   stop_

save_


    ##########################
    #  Experimental source  #
    ##########################


save_experimental_source
   _Saveframe_category   experimental_source


   loop_
      _Mol_label
      _Production_method
      _Host_organism_name_common
      _Genus
      _Species
      _Strain
      _Vector_name

      $CAP
      'recombinant technology'
       .
       .
       .
       .
       .


   stop_
```

```
save_


####################################
#  Sample contents and methodology  #
####################################

    ########################
    #  Sample description  #
    ########################


save_sample_1
   _Saveframe_category   sample

   _Sample_type          solution

   loop_
      _Mol_label
      _Concentration_value
      _Concentration_value_units
      _Concentration_min_value
      _Concentration_max_value
      _Isotopic_labeling

      $CAP  .  mM  0.8  1.2 "[U-15N]"

   stop_

save_


save_sample_2
   _Saveframe_category   sample

   _Sample_type          solution

   loop_
      _Mol_label
      _Concentration_value
      _Concentration_value_units
      _Isotopic_labeling

      $CAP  0.9  mM "[U-13C; U-15N]"

   stop_

save_


############################
#  Computer software used  #
############################


save_XWINNMR
   _Saveframe_category   software

   _Name                 XWINNMR

save_


save_Sparky
   _Saveframe_category   software

   _Name                 Sparky

save_


########################
#  Experimental detail  #
########################

    ###################################
    #  NMR Spectrometer definitions  #
    ###################################


save_NMR_spectrometer_1
   _Saveframe_category   NMR_spectrometer

   _Manufacturer         Bruker
   _Model                DRX
   _Field_strength       600

save_


save_NMR_spectrometer_2
   _Saveframe_category   NMR_spectrometer

   _Manufacturer         Bruker
   _Model                DMX
   _Field_strength       750

save_


    ############################
    #  NMR applied experiments  #
    ############################


save_NMR_applied_experiment
   _Saveframe_category   NMR_applied_experiment

   _Experiment_name
;
2D 1H-1H NOESY
2D 1H-13C HSQC
2D 1H-15N HSQC
3D 1H-1H-15N NOESY
3D 13C-1H-1H NOESY
3D CBCA(CO)NH
3D HNCA
3D HNCO
;

save_


######################
#  Sample conditions  #
######################


save_Exp-cond
   _Saveframe_category   sample_conditions


   loop_
```

```
      _Variable_type
      _Variable_value
      _Variable_value_error
      _Variable_value_units

       pH             7.3  0.1  n/a
       temperature  300    0.1  K

   stop_

save_


####################
#  NMR parameters  #
####################

    ##############################
    #  Assigned chemical shifts  #
    ##############################

#################################
#  Chemical shift referencing  #
#################################


save_chemical_shift_reference
   _Saveframe_category   chemical_shift_reference


   loop_
      _Mol_common_name
      _Atom_type
      _Atom_isotope_number
      _Atom_group
      _Chem_shift_units
      _Chem_shift_value
      _Reference_method
      _Reference_type
      _External_reference_sample_geometry
      _External_reference_location
      _External_reference_axis
      _Indirect_shift_ratio

TSP  H   1   methyl           ppm   0.00  .  direct .    1.0
TSP  C  13   methyl           ppm  -0.2   .  direct .    1.0
DSS  N  15  'methyl protons'  ppm   0.0   .  indirect    0.101329118

   stop_

save_


###################################
#  Assigned chemical shift lists  #
###################################

#####################################################################
#       Chemical Shift Ambiguity Index Value Definitions            #
#                                                                   #
#   Index Value            Definition                             #
#                                                                   #
#      1            Unique                                        #
#      2            Ambiguity of geminal atoms or geminal methyl #
#                        proton groups                            #
#      3            Aromatic atoms on opposite sides of the ring #
#                       (e.g. Tyr HE1 and HE2 protons)            #
#      4            Intraresidue ambiguities (e.g. Lys HG and    #
#                        HD protons)                              #
#      5            Interresidue ambiguities (Lys 12 vs. Lys 27) #
#      9            Ambiguous, specific ambiguity not defined    #
#                                                                   #
#####################################################################


save_shift_set_1
   _Saveframe_category               assigned_chemical_shifts


   loop_
      _Sample_label

      $sample_1
      $sample_2

   stop_

   _Sample_conditions_label         $Exp-cond
   _Chem_shift_reference_set_label  $chemical_shift_reference
   _Mol_system_component_name       "CAP N-terminus"

   loop_
      _Atom_shift_assign_ID
      _Residue_seq_code
      _Residue_label
      _Atom_name
      _Atom_type
      _Chem_shift_value
      _Chem_shift_value_error
      _Chem_shift_ambiguity_code

    1     3   LYS  C     C  178.355     0.1   1
    2     3   LYS  CA    C   59.337     0.1   1
    3     3   LYS  CB    C   28.034     0.3   1
    4     4   GLU  H     H    8.034     0.02  1
    5     4   GLU  C     C  176.826     0.1   1
    6     4   GLU  CA    C   61.952     0.1   1
    7     4   GLU  N     N  123.843     0.05  1
    8     5   PHE  H     H    8.489     0.02  1
    9     5   PHE  C     C  176.00      0.1   1
   10     5   PHE  CA    C   58.111     0.1   1
   11     5   PHE  N     N  119.60      0.05  1
   12     6   GLN  H     H    8.167     0.02  1
   13     6   GLN  C     C  178.320     0.1   1
   14     6   GLN  CA    C   59.281     0.1   1
   15     6   GLN  CB    C   32.00      0.3   1
   16     6   GLN  N     N  118.525     0.05  1
   17     7   ASN  H     H    8.169     0.02  1
   18     7   ASN  C     C  177.530     0.1   1
   19     7   ASN  CA    C   56.234     0.1   1
   20     7   ASN  CB    C   37.636     0.3   1
   21     7   ASN  N     N  118.851     0.05  1
   22     8   LEU  H     H    7.354     0.02  1
   23     8   LEU  N     N  120.373     0.05  1
   24     8   LEU  C     C  178.512     0.1   1
   25     8   LEU  CA    C   58.004     0.1   1
   26     8   LEU  CB    C   42.148     0.3   1
   27     9   VAL  H     H    7.397     0.02  1
   28     9   VAL  HA    H    3.389     0.05  1
   29     9   VAL  HB    H    2.293     0.05  1
   30     9   VAL  HG1   H    0.889     0.05  2
   31     9   VAL  HG2   H    0.723     0.05  2
   32     9   VAL  C     C  178.009     0.1   1
   33     9   VAL  CA    C   66.283     0.1   1
```

```
  34    9  VAL  CB   C   31.320    0.3   1
  35    9  VAL  N    N  122.06     0.05  1
  36   10  ASP  H    H    8.749    0.02  1
  37   10  ASP  HA   H    4.348    0.05  1
  38   10  ASP  HB2  H    2.736    0.05  2
  39   10  ASP  HB3  H    2.692    0.05  2
  40   10  ASP  C    C  177.654    0.1   1
  41   10  ASP  CA   C   57.309    0.1   1
  42   10  ASP  CB   C   40.517    0.3   1
  43   10  ASP  N    N  121.254    0.05  1
  44   11  GLN  H    H    7.781    0.02  1
  45   11  GLN  HA   H    3.881    0.05  1
  46   11  GLN  HB2  H    1.846    0.05  2
  47   11  GLN  HB3  H    1.468    0.05  2
  48   11  GLN  C    C  176.592    0.1   1
  49   11  GLN  CA   C   58.530    0.1   1
  50   11  GLN  CB   C   28.957    0.3   1
  51   11  GLN  N    N  116.680    0.05  1
  52   12  HIS  H    H    7.480    0.02  1
  53   12  HIS  HA   H    4.959    0.05  1
  54   12  HIS  HB2  H    3.359    0.05  2
  55   12  HIS  HB3  H    2.995    0.05  2
  56   12  HIS  C    C  176.668    0.1   1
  57   12  HIS  CA   C   56.908    0.1   1
  58   12  HIS  CB   C   35.027    0.3   1
  59   12  HIS  N    N  112.447    0.05  1
  60   13  ILE  H    H    8.172    0.02  1
  61   13  ILE  HA   H    4.095    0.05  1
  62   13  ILE  HB   H    2.694    0.05  1
  63   13  ILE  HG12 H    1.189    0.05  2
  64   13  ILE  HG13 H    0.579    0.05  2
  65   13  ILE  HG2  H    0.838    0.05  1
  66   13  ILE  HD1  H   -0.289    0.05  1
  67   13  ILE  C    C  177.415    0.1   1
  68   13  ILE  CA   C   59.928    0.1   1
  69   13  ILE  CB   C   33.81     0.3   1
  70   13  ILE  CG1  C   27.638    0.1   1
  71   13  ILE  CG2  C   17.846    0.1   1
  72   13  ILE  CD1  C    5.715    0.1   1
  73   13  ILE  N    N  119.564    0.05  1
  74   14  THR  H    H    8.023    0.02  1
  75   14  THR  HA   H    4.091    0.02  1
  76   14  THR  HB   H    4.516    0.02  1
  77   14  THR  HG2  H    1.281    0.05  1
  78   14  THR  CA   C   69.387    0.1   1
  79   14  THR  CB   C   66.214    0.1   1
  80   14  THR  CG2  C   22.595    0.1   1
  81   14  THR  N    N  119.258    0.05  1
  82   15  PRO  HA   H    4.509    0.02  1
  83   15  PRO  HB2  H    2.339    0.02  1
  84   15  PRO  HB3  H    1.938    0.02  1
  85   15  PRO  C    C  178.322    0.1   1
  86   15  PRO  CA   C   65.484    0.1   1
  87   15  PRO  CB   C   30.601    0.3   1
  88   16  PHE  H    H    6.969    0.02  1
  89   16  PHE  HA   H    4.411    0.05  1
  90   16  PHE  HB2  H    3.487    0.05  2
  91   16  PHE  HB3  H    3.302    0.05  2
  92   16  PHE  HD1  H    6.858    0.05  3
  93   16  PHE  HE1  H    5.755    0.05  3
  94   16  PHE  HZ   H    5.949    0.05  1
  95   16  PHE  C    C  178.815    0.1   1
  96   16  PHE  CA   C   60.711    0.1   1
  97   16  PHE  CB   C   40.022    0.3   1
  98   16  PHE  N    N  121.835    0.05  1
  99   17  VAL  H    H    8.895    0.02  1
 100   17  VAL  HA   H    3.395    0.05  1
 101   17  VAL  HB   H    2.221    0.05  1
 102   17  VAL  HG1  H    1.078    0.05  2
 103   17  VAL  HG2  H    0.975    0.05  2
 104   17  VAL  C    C  176.905    0.1   1
 105   17  VAL  CA   C   67.007    0.1   1
 106   17  VAL  CB   C   31.664    0.3   1
 107   17  VAL  N    N  126.060    0.05  1
 108   18  ALA  H    H    6.876    0.02  1
 109   18  ALA  HA   H    4.088    0.05  1
 110   18  ALA  HB   H    1.517    0.05  1
 111   18  ALA  C    C  175.548    0.1   1
 112   18  ALA  CA   C   55.363    0.1   1
 113   18  ALA  CB   C   17.691    0.3   1
 114   18  ALA  N    N  122.452    0.05  1
 115   19  LEU  H    H    7.964    0.02  1
 116   19  LEU  HA   H    4.045    0.05  1
 117   19  LEU  HB2  H    1.642    0.05  2
 118   19  LEU  HB3  H    2.048    0.05  2
 119   19  LEU  HG   H    1.540    0.05  1
 120   19  LEU  HD1  H    1.083    0.05  2
 121   19  LEU  HD2  H    1.167    0.05  2
 122   19  LEU  C    C  177.882    0.1   1
 123   19  LEU  CA   C   57.692    0.1   1
 124   19  LEU  CB   C   43.611    0.3   1
 125   19  LEU  N    N  119.362    0.05  1
 126   20  SER  H    H    8.084    0.02  1
 127   20  SER  HA   H    4.115    0.05  1
 128   20  SER  HB2  H    3.907    0.05  2
 129   20  SER  HB3  H    3.416    0.05  2
 130   20  SER  C    C  174.415    0.1   1
 131   20  SER  CA   C   63.514    0.1   1
 132   20  SER  N    N  116.413    0.05  1
 133   21  LYS  H    H    7.487    0.02  1
 134   21  LYS  HA   H    3.721    0.05  1
 135   21  LYS  HB2  H    1.840    0.05  2
 136   21  LYS  HB3  H    1.634    0.05  2
 137   21  LYS  HG3  H    1.377    0.05  2
 138   21  LYS  C    C  177.538    0.1   1
 139   21  LYS  CA   C   58.844    0.1   1
 140   21  LYS  CB   C   32.309    0.3   1
 141   21  LYS  N    N  117.655    0.05  1
 142   22  LYS  H    H    7.210    0.02  1
 143   22  LYS  HA   H    3.975    0.05  1
 144   22  LYS  HB2  H    1.904    0.05  2
 145   22  LYS  HB3  H    1.725    0.05  2
 146   22  LYS  HG2  H    1.379    0.05  2
 147   22  LYS  HG3  H    1.523    0.05  2
 148   22  LYS  C    C  177.654    0.1   1
 149   22  LYS  CA   C   57.947    0.1   1
 150   22  LYS  CB   C   32.867    0.3   1
 151   22  LYS  N    N  118.722    0.05  1
 152   23  LEU  H    H    7.364    0.02  1
 153   23  LEU  HA   H    3.799    0.05  1
 154   23  LEU  HB2  H    1.588    0.05  2
 155   23  LEU  HB3  H    1.396    0.05  2
 156   23  LEU  HG   H    0.674    0.05  1
 157   23  LEU  HD1  H    0.492    0.05  2
 158   23  LEU  HD2  H    0.139    0.05  2
 159   23  LEU  C    C  176.250    0.1   1
 160   23  LEU  CA   C   57.366    0.1   1
 161   23  LEU  CB   C   41.935    0.3   1
 162   23  LEU  CG   C   25.97     0.1   1
 163   23  LEU  CD1  C   22.799    0.1   2
 164   23  LEU  CD2  C   25.302    0.1   2
 165   23  LEU  N    N  120.072    0.05  1
 166   24  ALA  H    H    6.766    0.02  1
 167   24  ALA  HA   H    4.331    0.05  1
```

168	24	ALA	HB	H	0.579	0.05	1
169	24	ALA	CA	C	50.754	0.1	1
170	24	ALA	CB	C	19.990	0.3	1
171	24	ALA	N	N	115.055	0.05	1
172	25	PRO	HD2	H	3.766	0.05	2
173	25	PRO	C	C	178.468	0.1	1
174	25	PRO	CA	C	65.067	0.1	1
175	25	PRO	CB	C	31.772	0.3	1
176	26	GLU	H	H	10.698	0.02	1
177	26	GLU	HA	H	4.277	0.05	1
178	26	GLU	HB2	H	1.789	0.05	2
179	26	GLU	HB3	H	2.221	0.05	2
180	26	GLU	HG2	H	2.899	0.05	2
181	26	GLU	C	C	177.504	0.1	1
182	26	GLU	CA	C	58.944	0.1	1
183	26	GLU	CB	C	27.066	0.3	1
184	26	GLU	N	N	118.558	0.05	1
185	27	VAL	H	H	6.476	0.02	1
186	27	VAL	HA	H	3.668	0.05	1
187	27	VAL	HB	H	2.011	0.05	1
188	27	VAL	HG1	H	0.649	0.05	2
189	27	VAL	HG2	H	0.949	0.05	2
190	27	VAL	C	C	177.554	0.1	1
191	27	VAL	CA	C	65.354	0.1	1
192	27	VAL	CB	C	30.525	0.3	1
193	27	VAL	CG1	C	23.468	0.1	2
194	27	VAL	CG2	C	21.231	0.1	2
195	27	VAL	N	N	121.888	0.05	1
196	28	GLY	H	H	8.016	0.02	1
197	28	GLY	HA2	H	3.861	0.05	2
198	28	GLY	HA3	H	3.284	0.05	2
199	28	GLY	C	C	175.662	0.1	1
200	28	GLY	CA	C	47.808	0.1	1
201	28	GLY	N	N	108.085	0.05	1
202	29	ASN	H	H	8.710	0.02	1
203	29	ASN	HA	H	4.501	0.05	1
204	29	ASN	HB2	H	2.926	0.05	2
205	29	ASN	HB3	H	2.843	0.05	2
206	29	ASN	HD21	H	7.557	0.05	2
207	29	ASN	C	C	177.694	0.1	1
208	29	ASN	CA	C	55.678	0.1	1
209	29	ASN	CB	C	38.024	0.3	1
210	29	ASN	N	N	120.106	0.05	1
211	29	ASN	ND2	N	111.968	0.05	1
212	30	GLN	H	H	7.909	0.02	1
213	30	GLN	HA	H	3.953	0.05	1
214	30	GLN	HB2	H	2.171	0.05	2
215	30	GLN	HB3	H	2.617	0.05	2
216	30	GLN	HG2	H	2.456	0.05	2
217	30	GLN	C	C	178.382	0.1	1
218	30	GLN	CA	C	59.889	0.1	1
219	30	GLN	CB	C	29.376	0.3	1
220	30	GLN	N	N	121.957	0.05	1
221	31	VAL	H	H	8.614	0.02	1
222	31	VAL	HA	H	3.784	0.05	1
223	31	VAL	HB	H	2.180	0.05	1
224	31	VAL	HG1	H	1.186	0.05	2
225	31	VAL	HG2	H	0.972	0.05	2
226	31	VAL	C	C	178.242	0.1	1
227	31	VAL	CA	C	66.570	0.1	1
228	31	VAL	CB	C	31.653	0.3	1
229	31	VAL	CG1	C	24.727	0.1	2
230	31	VAL	N	N	120.492	0.05	1
231	32	GLU	H	H	8.355	0.02	1
232	32	GLU	HA	H	4.491	0.05	1
233	32	GLU	HB2	H	2.151	0.05	2
234	32	GLU	HG2	H	2.349	0.05	2
235	32	GLU	C	C	178.990	0.1	1
236	32	GLU	CA	C	60.274	0.1	1
237	32	GLU	CB	C	29.311	0.3	1
238	32	GLU	N	N	122.738	0.05	1
239	33	GLN	H	H	7.154	0.02	1
240	33	GLN	HA	H	4.024	0.05	1
241	33	GLN	HB2	H	2.261	0.05	2
242	33	GLN	HG2	H	2.715	0.05	2
243	33	GLN	HG3	H	2.483	0.05	2
244	33	GLN	C	C	177.396	0.1	1
245	33	GLN	CA	C	57.560	0.1	1
246	33	GLN	CB	C	27.38	0.3	1
247	33	GLN	N	N	116.783	0.05	1
248	34	LEU	H	H	7.578	0.02	1
249	34	LEU	HA	H	3.980	0.05	1
250	34	LEU	HB2	H	1.728	0.05	2
251	34	LEU	HG	H	1.441	0.05	1
252	34	LEU	C	C	177.615	0.1	1
253	34	LEU	CA	C	58.372	0.1	1
254	34	LEU	CB	C	41.752	0.3	1
255	34	LEU	N	N	121.641	0.05	1
256	35	VAL	H	H	8.732	0.02	1
257	35	VAL	HA	H	3.438	0.05	1
258	35	VAL	HB	H	2.198	0.05	1
259	35	VAL	HG1	H	0.989	0.05	2
260	35	VAL	HG2	H	0.898	0.05	2
261	35	VAL	C	C	177.677	0.1	1
262	35	VAL	CA	C	67.421	0.1	1
263	35	VAL	CB	C	31.009	0.3	1
264	35	VAL	N	N	120.367	0.05	1
265	36	LYS	H	H	7.229	0.02	1
266	36	LYS	HA	H	3.836	0.05	1
267	36	LYS	HB2	H	0.554	0.05	2
268	36	LYS	HB3	H	1.144	0.05	2
269	36	LYS	HG2	H	1.540	0.05	2
270	36	LYS	HG3	H	1.282	0.05	2
271	36	LYS	HD2	H	1.275	0.05	2
272	36	LYS	HE2	H	2.659	0.05	2
273	36	LYS	C	C	179.008	0.1	1
274	36	LYS	CA	C	59.644	0.1	1
275	36	LYS	CB	C	30.729	0.3	1
276	36	LYS	N	N	120.521	0.05	1
277	37	ALA	H	H	7.740	0.02	1
278	37	ALA	HA	H	4.111	0.05	1
279	37	ALA	HB	H	1.722	0.05	1
280	37	ALA	C	C	178.454	0.1	1
281	37	ALA	CA	C	55.237	0.1	1
282	37	ALA	CB	C	18.654	0.3	1
283	37	ALA	N	N	123.933	0.05	1
284	38	ILE	H	H	8.668	0.02	1
285	38	ILE	HA	H	3.752	0.05	1
286	38	ILE	HB	H	2.188	0.05	1
287	38	ILE	HG2	H	1.190	0.05	1
288	38	ILE	HD1	H	0.845	0.05	1
289	38	ILE	C	C	177.720	0.1	1
290	38	ILE	CA	C	65.919	0.1	1
291	38	ILE	CB	C	37.519	0.3	1
292	38	ILE	N	N	121.456	0.05	1
293	39	ASP	H	H	8.529	0.02	1
294	39	ASP	HA	H	4.641	0.05	1
295	39	ASP	HB2	H	2.689	0.05	2
296	39	ASP	HB3	H	2.835	0.05	2
297	39	ASP	C	C	179.029	0.1	1
298	39	ASP	CA	C	57.664	0.1	1
299	39	ASP	CB	C	40.388	0.3	1
300	39	ASP	N	N	122.977	0.05	1
301	40	ALA	H	H	8.040	0.02	1

302	40	ALA	HA	H	4.349	0.05	1
303	40	ALA	HB	H	1.774	0.05	1
304	40	ALA	C	C	179.719	0.1	1
305	40	ALA	CA	C	54.793	0.1	1
306	40	ALA	CB	C	17.032	0.3	1
307	40	ALA	N	N	125.092	0.05	1
308	41	GLU	H	H	8.749	0.02	1
309	41	GLU	HA	H	3.929	0.05	1
310	41	GLU	HB2	H	2.079	0.05	2
311	41	GLU	HG2	H	2.781	0.05	2
312	41	GLU	C	C	176.689	0.1	1
313	41	GLU	CA	C	60.547	0.1	1
314	41	GLU	CB	C	28.097	0.3	1
315	41	GLU	N	N	123.601	0.05	1
316	42	LYS	H	H	8.725	0.02	1
317	42	LYS	HA	H	3.395	0.05	1
318	42	LYS	HB2	H	2.158	0.05	2
319	42	LYS	HB3	H	1.562	0.05	2
320	42	LYS	HG2	H	1.568	0.05	2
321	42	LYS	HE3	H	2.384	0.05	2
322	42	LYS	C	C	177.557	0.1	1
323	42	LYS	CA	C	60.387	0.1	1
324	42	LYS	CB	C	31.729	0.3	1
325	42	LYS	N	N	122.564	0.05	1
326	43	ALA	H	H	7.720	0.02	1
327	43	ALA	HA	H	4.196	0.05	1
328	43	ALA	HB	H	1.629	0.05	1
329	43	ALA	C	C	179.469	0.1	1
330	43	ALA	CA	C	54.847	0.1	1
331	43	ALA	CB	C	17.483	0.3	1
332	43	ALA	N	N	120.920	0.05	1
333	44	LEU	H	H	7.716	0.02	1
334	44	LEU	HA	H	4.091	0.05	1
335	44	LEU	HB2	H	1.626	0.05	2
336	44	LEU	HB3	H	1.845	0.05	2
337	44	LEU	HG	H	1.423	0.05	1
338	44	LEU	HD1	H	0.721	0.05	2
339	44	LEU	HD2	H	0.930	0.05	2
340	44	LEU	C	C	177.614	0.1	1
341	44	LEU	CA	C	57.926	0.1	1
342	44	LEU	CB	C	41.376	0.3	1
343	44	LEU	CD1	C	26.189	0.1	2
344	44	LEU	N	N	122.384	0.05	1
345	45	ILE	H	H	8.091	0.02	1
346	45	ILE	HA	H	3.289	0.05	1
347	45	ILE	HB	H	1.514	0.05	1
348	45	ILE	HG12	H	1.773	0.05	2
349	45	ILE	HG13	H	0.930	0.05	2
350	45	ILE	HG2	H	0.539	0.05	1
351	45	ILE	HD1	H	0.327	0.05	1
352	45	ILE	C	C	178.470	0.1	1
353	45	ILE	CA	C	66.045	0.1	1
354	45	ILE	CB	C	38.110	0.3	1
355	45	ILE	CG1	C	16.767	0.1	2
356	45	ILE	CD1	C	14.119	0.1	1
357	45	ILE	N	N	121.725	0.05	1
358	46	ASN	H	H	8.238	0.02	1
359	46	ASN	HA	H	4.346	0.05	1
360	46	ASN	HB2	H	2.733	0.05	2
361	46	ASN	HB3	H	3.179	0.05	2
362	46	ASN	C	C	178.265	0.1	1
363	46	ASN	CA	C	56.948	0.1	1
364	46	ASN	CB	C	39.034	0.3	1
365	46	ASN	N	N	119.760	0.05	1
366	47	THR	H	H	8.989	0.02	1
367	47	THR	HA	H	3.832	0.05	1
368	47	THR	HB	H	4.400	0.05	1
369	47	THR	HG2	H	1.374	0.05	1
370	47	THR	C	C	177.124	0.1	1
371	47	THR	CA	C	67.399	0.1	1
372	47	THR	CB	C	69.287	0.3	1
373	47	THR	N	N	122.728	0.05	1
374	48	ALA	H	H	8.646	0.02	1
375	48	ALA	HA	H	3.998	0.05	1
376	48	ALA	HB	H	0.007	0.05	1
377	48	ALA	C	C	177.406	0.1	1
378	48	ALA	CA	C	55.438	0.1	1
379	48	ALA	CB	C	15.463	0.3	1
380	48	ALA	N	N	128.647	0.05	1
381	49	SER	H	H	7.593	0.02	1
382	49	SER	HA	H	4.312	0.05	1
383	49	SER	HB2	H	4.138	0.05	2
384	49	SER	C	C	175.645	0.1	1
385	49	SER	CA	C	60.907	0.1	1
386	49	SER	CB	C	63.486	0.3	1
387	49	SER	N	N	111.216	0.05	1
388	50	GLN	H	H	7.502	0.02	1
389	50	GLN	HA	H	4.831	0.05	1
390	50	GLN	HB2	H	2.105	0.05	2
391	50	GLN	HG2	H	2.473	0.05	2
392	50	GLN	C	C	176.312	0.1	1
393	50	GLN	CA	C	55.071	0.1	1
394	50	GLN	CB	C	32.148	0.3	1
395	50	GLN	N	N	118.997	0.05	1
396	51	SER	H	H	7.749	0.02	1
397	51	SER	HA	H	5.344	0.05	1
398	51	SER	HB2	H	4.068	0.05	2
399	51	SER	HB3	H	3.912	0.05	2
400	51	SER	C	C	172.390	0.1	1
401	51	SER	CA	C	58.328	0.1	1
402	51	SER	CB	C	66.612	0.3	1
403	51	SER	N	N	116.870	0.05	1
404	52	LYS	H	H	8.464	0.02	1
405	52	LYS	HA	H	4.386	0.05	1
406	52	LYS	HG2	H	1.658	0.05	2
407	52	LYS	HG3	H	1.352	0.05	2
408	52	LYS	C	C	176.781	0.1	1
409	52	LYS	CA	C	55.968	0.1	1
410	52	LYS	CB	C	33.190	0.3	1
411	52	LYS	N	N	124.416	0.05	1
412	53	LYS	H	H	8.654	0.02	1
413	53	LYS	HA	H	3.050	0.05	1
414	53	LYS	HB2	H	1.793	0.05	2
415	53	LYS	HB3	H	2.427	0.05	2
416	53	LYS	HG2	H	1.112	0.05	2
417	53	LYS	CA	C	54.745	0.1	1
418	53	LYS	CG	C	24.152	0.1	1
419	53	LYS	N	N	126.401	0.05	1
420	54	PRO	C	C	175.630	0.1	1
421	54	PRO	CA	C	61.669	0.1	1
422	54	PRO	CB	C	32.244	0.3	1
423	55	SER	H	H	8.259	0.02	1
424	55	SER	HA	H	4.278	0.05	1
425	55	SER	HB2	H	4.184	0.05	2
426	55	SER	HB3	H	4.046	0.05	2
427	55	SER	C	C	174.948	0.1	1
428	55	SER	CA	C	57.501	0.1	1
429	55	SER	CB	C	64.581	0.3	1
430	55	SER	N	N	114.111	0.05	1
431	56	GLN	H	H	8.800	0.02	1
432	56	GLN	HA	H	3.939	0.05	1
433	56	GLN	HB2	H	2.174	0.05	2
434	56	GLN	HB3	H	2.025	0.05	2
435	56	GLN	HG2	H	2.454	0.05	2

436	56	GLN	C	C	177.519	0.1	1
437	56	GLN	CA	C	59.362	0.1	1
438	56	GLN	CB	C	28.076	0.3	1
439	56	GLN	N	N	122.598	0.05	1
440	57	GLU	H	H	8.749	0.02	1
441	57	GLU	HA	H	3.974	0.05	1
442	57	GLU	HB2	H	2.061	0.05	2
443	57	GLU	C	C	178.536	0.1	1
444	57	GLU	CA	C	60.325	0.1	1
445	57	GLU	CB	C	28.699	0.3	1
446	57	GLU	N	N	118.276	0.05	1
447	58	THR	H	H	7.650	0.02	1
448	58	THR	HA	H	4.025	0.05	1
449	58	THR	HB	H	4.145	0.05	1
450	58	THR	HG2	H	1.204	0.05	1
451	58	THR	C	C	176.230	0.1	1
452	58	THR	CA	C	66.162	0.1	1
453	58	THR	CB	C	67.912	0.3	1
454	58	THR	N	N	119.528	0.05	1
455	59	LEU	H	H	8.427	0.02	1
456	59	LEU	HA	H	3.867	0.05	1
457	59	LEU	HB2	H	2.193	0.05	2
458	59	LEU	HB3	H	1.868	0.05	2
459	59	LEU	HG	H	1.471	0.05	1
460	59	LEU	HD2	H	0.953	0.05	2
461	59	LEU	C	C	177.613	0.1	1
462	59	LEU	CA	C	58.759	0.1	1
463	59	LEU	CB	C	41.440	0.3	1
464	59	LEU	N	N	125.158	0.05	1
465	60	LEU	H	H	7.876	0.02	1
466	60	LEU	HA	H	3.995	0.05	1
467	60	LEU	HB2	H	1.904	0.05	2
468	60	LEU	HB3	H	1.511	0.05	2
469	60	LEU	HD2	H	0.934	0.05	2
470	60	LEU	C	C	179.125	0.1	1
471	60	LEU	CA	C	57.604	0.1	1
472	60	LEU	CB	C	41.097	0.3	1
473	60	LEU	CG	C	26.233	0.1	1
474	60	LEU	N	N	116.841	0.05	1
475	61	GLU	H	H	7.264	0.02	1
476	61	GLU	HA	H	4.100	0.05	1
477	61	GLU	HB2	H	2.221	0.05	2
478	61	GLU	HB3	H	1.880	0.05	2
479	61	GLU	HG2	H	2.246	0.05	2
480	61	GLU	HG3	H	2.334	0.05	2
481	61	GLU	C	C	179.347	0.1	1
482	61	GLU	CA	C	58.886	0.1	1
483	61	GLU	CB	C	29.419	0.3	1
484	61	GLU	N	N	118.939	0.05	1
485	62	LEU	H	H	8.209	0.02	1
486	62	LEU	HA	H	4.161	0.05	1
487	62	LEU	HB2	H	1.540	0.05	2
488	62	LEU	HB3	H	2.171	0.05	2
489	62	LEU	HG	H	2.007	0.05	1
490	62	LEU	HD2	H	0.933	0.05	2
491	62	LEU	C	C	178.487	0.1	1
492	62	LEU	CA	C	57.313	0.1	1
493	62	LEU	CB	C	41.902	0.3	1
494	62	LEU	CG	C	26.422	0.1	1
495	62	LEU	N	N	120.771	0.05	1
496	63	ILE	H	H	7.515	0.02	1
497	63	ILE	HA	H	4.420	0.05	1
498	63	ILE	HB	H	2.127	0.05	1
499	63	ILE	HG12	H	1.584	0.05	2
500	63	ILE	HG13	H	1.329	0.05	2
501	63	ILE	HG2	H	0.931	0.05	1
502	63	ILE	HD1	H	0.848	0.05	1
503	63	ILE	C	C	175.706	0.1	1
504	63	ILE	CA	C	61.152	0.1	1
505	63	ILE	CB	C	38.271	0.3	1
506	63	ILE	CD1	C	17.473	0.1	1
507	63	ILE	N	N	109.825	0.05	1
508	64	LYS	H	H	7.321	0.02	1
509	64	LYS	HA	H	4.126	0.05	1
510	64	LYS	HB2	H	2.053	0.05	2
511	64	LYS	HB3	H	1.903	0.05	2
512	64	LYS	HG2	H	1.435	0.05	2
513	64	LYS	CA	C	61.830	0.1	1
514	64	LYS	N	N	124.685	0.05	1
515	65	PRO	HA	H	4.116	0.05	1
516	65	PRO	HB2	H	2.377	0.05	2
517	65	PRO	HB3	H	1.861	0.05	2
518	65	PRO	C	C	178.141	0.1	1
519	65	PRO	CA	C	66.331	0.1	1
520	65	PRO	CB	C	30.686	0.3	1
521	66	LEU	H	H	7.312	0.02	1
522	66	LEU	HA	H	4.063	0.05	1
523	66	LEU	HB3	H	1.915	0.05	2
524	66	LEU	HG	H	1.749	0.05	1
525	66	LEU	HD1	H	0.900	0.05	2
526	66	LEU	C	C	178.011	0.1	1
527	66	LEU	CA	C	59.121	0.1	1
528	66	LEU	CB	C	41.602	0.3	1
529	66	LEU	N	N	117.001	0.05	1
530	67	ASN	H	H	7.930	0.02	1
531	67	ASN	HA	H	4.444	0.05	1
532	67	ASN	HB2	H	2.805	0.05	2
533	67	ASN	HB3	H	2.946	0.05	2
534	67	ASN	HD21	H	7.502	0.05	2
535	67	ASN	HD22	H	6.417	0.05	2
536	67	ASN	C	C	177.005	0.1	1
537	67	ASN	CA	C	55.293	0.1	1
538	67	ASN	CB	C	37.476	0.3	1
539	67	ASN	N	N	118.238	0.05	1
540	67	ASN	ND2	N	110.380	0.05	1
541	68	ASN	H	H	8.372	0.02	1
542	68	ASN	HA	H	4.481	0.05	1
543	68	ASN	HB2	H	2.670	0.05	2
544	68	ASN	C	C	177.994	0.1	1
545	68	ASN	CA	C	56.103	0.1	1
546	68	ASN	CB	C	37.498	0.3	1
547	68	ASN	N	N	122.745	0.05	1
548	69	PHE	H	H	7.988	0.02	1
549	69	PHE	HA	H	4.655	0.05	1
550	69	PHE	HB2	H	2.922	0.05	2
551	69	PHE	HB3	H	3.442	0.05	2
552	69	PHE	C	C	177.973	0.1	1
553	69	PHE	CA	C	61.358	0.1	1
554	69	PHE	CB	C	39.109	0.3	1
555	69	PHE	N	N	120.956	0.05	1
556	70	ALA	H	H	8.346	0.02	1
557	70	ALA	HA	H	3.948	0.05	1
558	70	ALA	HB	H	1.526	0.05	1
559	70	ALA	C	C	179.701	0.1	1
560	70	ALA	CA	C	55.594	0.1	1
561	70	ALA	CB	C	18.375	0.3	1
562	70	ALA	N	N	121.735	0.05	1
563	71	ALA	H	H	8.121	0.02	1
564	71	ALA	HA	H	3.953	0.05	1
565	71	ALA	HB	H	1.633	0.05	1
566	71	ALA	C	C	179.984	0.1	1
567	71	ALA	CA	C	54.817	0.1	1
568	71	ALA	CB	C	17.687	0.3	1
569	71	ALA	N	N	122.028	0.05	1

```
570   72   GLU  H     H    7.793     0.02  1
571   72   GLU  HA    H    4.197     0.05  1
572   72   GLU  HB2   H    2.011     0.05  2
573   72   GLU  HG2   H    2.431     0.05  2
574   72   GLU  HG3   H    2.279     0.05  2
575   72   GLU  C     C  178.481     0.1   1
576   72   GLU  CA    C   58.905     0.1   1
577   72   GLU  CB    C   29.011     0.3   1
578   72   GLU  N     N  121.288     0.05  1
579   73   VAL  H     H    8.006     0.02  1
580   73   VAL  HA    H    3.681     0.05  1
581   73   VAL  HB    H    2.690     0.05  1
582   73   VAL  HG1   H    0.961     0.05  2
583   73   VAL  HG2   H    1.190     0.05  2
584   73   VAL  C     C  178.661     0.1   1
585   73   VAL  CA    C   67.274     0.1   1
586   73   VAL  CB    C   31.439     0.3   1
587   73   VAL  CG1   C   22.595     0.1   2
588   73   VAL  N     N  119.742     0.05  1
589   74   GLY  H     H    7.693     0.02  1
590   74   GLY  HA2   H    3.936     0.05  2
591   74   GLY  HA3   H    4.135     0.05  2
592   74   GLY  C     C  175.551     0.1   1
593   74   GLY  CA    C   47.465     0.1   1
594   74   GLY  N     N  108.236     0.05  1
595   75   LYS  H     H    8.160     0.02  1
596   75   LYS  HA    H    4.163     0.05  1
597   75   LYS  HB2   H    2.019     0.05  2
598   75   LYS  HG2   H    1.493     0.05  2
599   75   LYS  HD2   H    1.689     0.05  2
600   75   LYS  C     C  179.339     0.1   1
601   75   LYS  CA    C   59.703     0.1   1
602   75   LYS  CB    C   32.448     0.3   1
603   75   LYS  N     N  124.579     0.05  1
604   76   ILE  H     H    8.135     0.02  1
605   76   ILE  HA    H    3.754     0.05  1
606   76   ILE  HB    H    1.777     0.05  1
607   76   ILE  HG12  H    1.166     0.05  2
608   76   ILE  HG13  H    1.996     0.05  2
609   76   ILE  HG2   H    0.997     0.05  1
610   76   ILE  HD1   H    0.883     0.05  1
611   76   ILE  C     C  178.001     0.1   1
612   76   ILE  CA    C   65.319     0.1   1
613   76   ILE  CB    C   39.088     0.3   1
614   76   ILE  CG2   C   17.846     0.1   2
615   76   ILE  N     N  120.580     0.05  1
616   77   ARG  H     H    7.095     0.02  1
617   77   ARG  HA    H    4.020     0.05  1
618   77   ARG  HB2   H    2.128     0.05  2
619   77   ARG  HB3   H    2.288     0.05  2
620   77   ARG  C     C  177.698     0.1   1
621   77   ARG  CA    C   60.897     0.1   1
622   77   ARG  CB    C   29.698     0.3   1
623   77   ARG  N     N  121.576     0.05  1
624   78   ASP  H     H    8.454     0.02  1
625   78   ASP  HA    H    4.466     0.05  1
626   78   ASP  HB2   H    2.839     0.05  2
627   78   ASP  C     C  177.707     0.1   1
628   78   ASP  CA    C   57.554     0.1   1
629   78   ASP  CB    C   40.678     0.3   1
630   78   ASP  N     N  119.061     0.05  1
631   79   SER  H     H    7.941     0.02  1
632   79   SER  HA    H    4.601     0.05  1
633   79   SER  HB2   H    4.068     0.05  2
634   79   SER  C     C  174.444     0.1   1
635   79   SER  CA    C   59.356     0.1   1
636   79   SER  CB    C   64.012     0.3   1
637   79   SER  N     N  113.311     0.05  1
638   80   ASN  H     H    7.671     0.02  1
639   80   ASN  HA    H    5.285     0.05  1
640   80   ASN  HB2   H    2.630     0.05  2
641   80   ASN  HB3   H    3.004     0.05  2
642   80   ASN  C     C  180.234     0.1   1
643   80   ASN  CA    C   53.534     0.1   1
644   80   ASN  CB    C   40.785     0.3   1
645   80   ASN  N     N  123.048     0.05  1
646   81   ARG  H     H    7.973     0.02  1
647   81   ARG  HA    H    3.894     0.05  1
648   81   ARG  HG2   H    1.692     0.05  2
649   81   ARG  C     C  176.915     0.1   1
650   81   ARG  CA    C   59.247     0.1   1
651   81   ARG  CB    C   29.579     0.3   1
652   81   ARG  N     N  119.533     0.05  1
653   82   SER  H     H    8.206     0.02  1
654   82   SER  HA    H    4.478     0.05  1
655   82   SER  HB2   H    4.140     0.05  2
656   82   SER  C     C  174.662     0.1   1
657   82   SER  CA    C   57.760     0.1   1
658   82   SER  CB    C   63.185     0.3   1
659   82   SER  N     N  111.962     0.05  1
660   83   SER  H     H    7.433     0.02  1
661   83   SER  HA    H    4.379     0.05  1
662   83   SER  C     C  177.698     0.1   1
663   83   SER  CA    C   58.363     0.1   1
664   83   SER  CB    C   64.001     0.3   1
665   83   SER  N     N  116.689     0.05  1
666   84   LYS  H     H    9.428     0.02  1
667   84   LYS  HA    H    4.281     0.05  1
668   84   LYS  HB2   H    1.683     0.05  2
669   84   LYS  HG2   H    1.420     0.05  2
670   84   LYS  HG3   H    0.849     0.05  2
671   84   LYS  C     C  176.661     0.1   1
672   84   LYS  CA    C   58.177     0.1   1
673   84   LYS  CB    C   31.267     0.3   1
674   84   LYS  N     N  131.639     0.05  1
675   85   PHE  H     H    9.161     0.02  1
676   85   PHE  HA    H    4.606     0.05  1
677   85   PHE  HB2   H    3.161     0.05  2
678   85   PHE  HB3   H    2.408     0.05  2
679   85   PHE  C     C  175.821     0.1   1
680   85   PHE  CA    C   56.983     0.1   1
681   85   PHE  CB    C   38.056     0.3   1
682   85   PHE  N     N  119.699     0.05  1
683   86   PHE  H     H    7.691     0.02  1
684   86   PHE  HA    H    4.188     0.05  1
685   86   PHE  HB2   H    3.133     0.05  2
686   86   PHE  HB3   H    3.059     0.05  2
687   86   PHE  C     C  178.256     0.1   1
688   86   PHE  CA    C   63.017     0.1   1
689   86   PHE  CB    C   39.131     0.3   1
690   86   PHE  N     N  124.183     0.05  1
691   87   ASN  H     H    9.777     0.02  1
692   87   ASN  HA    H    4.593     0.05  1
693   87   ASN  HB2   H    2.630     0.05  2
694   87   ASN  HD21  H    7.119     0.05  2
695   87   ASN  HD22  H    8.533     0.05  2
696   87   ASN  C     C  177.218     0.1   1
697   87   ASN  CA    C   56.867     0.1   1
698   87   ASN  CB    C   37.562     0.3   1
699   87   ASN  N     N  120.323     0.05  1
700   87   ASN  ND2   N  115.252     0.05  1
701   88   ASN  H     H    7.489     0.02  1
702   88   ASN  HA    H    4.149     0.05  1
703   88   ASN  HB2   H    1.936     0.05  2
```

704	88	ASN	HB3	H	2.759	0.05	2
705	88	ASN	C	C	175.929	0.1	1
706	88	ASN	CA	C	57.602	0.1	1
707	88	ASN	CB	C	38.336	0.3	1
708	88	ASN	N	N	119.005	0.05	1
709	89	LEU	H	H	7.870	0.02	1
710	89	LEU	HA	H	4.993	0.05	1
711	89	LEU	HB2	H	1.181	0.05	2
712	89	LEU	HB3	H	1.797	0.05	2
713	89	LEU	HD1	H	0.672	0.05	2
714	89	LEU	HD2	H	1.025	0.05	2
715	89	LEU	CA	C	55.691	0.1	1
716	89	LEU	CB	C	42.599	0.1	1
717	89	LEU	CD1	C	23.468	0.1	2
718	89	LEU	CD2	C	27.979	0.1	2
719	89	LEU	N	N	116.308	0.05	1
720	90	SER	C	C	174.488	0.1	1
721	90	SER	CA	C	60.380	0.1	1
722	90	SER	CB	C	61.649	0.3	1
723	91	ALA	H	H	7.940	0.02	1
724	91	ALA	HA	H	4.256	0.05	1
725	91	ALA	HB	H	1.844	0.05	1
726	91	ALA	C	C	178.298	0.1	1
727	91	ALA	CA	C	55.876	0.1	1
728	91	ALA	CB	C	18.074	0.3	1
729	91	ALA	N	N	125.059	0.05	1
730	92	ILE	H	H	6.904	0.02	1
731	92	ILE	HA	H	3.810	0.05	1
732	92	ILE	HB	H	2.147	0.05	1
733	92	ILE	HG12	H	1.517	0.05	2
734	92	ILE	HG13	H	1.164	0.05	2
735	92	ILE	HG2	H	0.879	0.05	1
736	92	ILE	HD1	H	0.506	0.05	1
737	92	ILE	C	C	177.902	0.1	1
738	92	ILE	CA	C	62.669	0.1	1
739	92	ILE	CB	C	36.391	0.3	1
740	92	ILE	CG1	C	26.801	0.1	2
741	92	ILE	CG2	C	17.846	0.1	2
742	92	ILE	CD1	C	10.728	0.1	1
743	92	ILE	N	N	113.418	0.05	1
744	93	SER	H	H	8.652	0.02	1
745	93	SER	HA	H	3.838	0.05	1
746	93	SER	HB2	H	4.107	0.05	2
747	93	SER	HB3	H	3.553	0.05	2
748	93	SER	C	C	177.569	0.1	1
749	93	SER	CA	C	63.732	0.1	1
750	93	SER	CB	C	62.025	0.3	1
751	93	SER	N	N	116.879	0.05	1
752	94	GLU	H	H	8.818	0.02	1
753	94	GLU	HA	H	4.422	0.05	2
754	94	GLU	HB2	H	2.081	0.05	2
755	94	GLU	HB3	H	2.479	0.05	2
756	94	GLU	HG2	H	2.404	0.05	2
757	94	GLU	C	C	176.329	0.1	1
758	94	GLU	CA	C	56.995	0.1	1
759	94	GLU	CB	C	28.097	0.3	1
760	94	GLU	N	N	118.807	0.05	1
761	95	SER	H	H	7.594	0.02	1
762	95	SER	HA	H	5.167	0.05	1
763	95	SER	HB2	H	4.374	0.05	2
764	95	SER	HB3	H	3.981	0.05	2
765	95	SER	C	C	176.192	0.1	1
766	95	SER	CA	C	57.358	0.1	1
767	95	SER	CB	C	67.611	0.3	1
768	95	SER	N	N	110.303	0.05	1
769	96	ILE	H	H	7.422	0.02	1
770	96	ILE	HA	H	4.526	0.05	1
771	96	ILE	HB	H	2.353	0.05	1
772	96	ILE	HG12	H	1.471	0.05	2
773	96	ILE	HG13	H	1.703	0.05	2
774	96	ILE	HG2	H	1.098	0.05	1
775	96	ILE	HD1	H	0.861	0.05	1
776	96	ILE	C	C	175.575	0.1	1
777	96	ILE	CA	C	60.151	0.1	1
778	96	ILE	CB	C	38.798	0.3	1
779	96	ILE	CG2	C	16.389	0.1	2
780	96	ILE	N	N	126.717	0.05	1
781	97	GLY	H	H	8.903	0.02	1
782	97	GLY	HA2	H	4.333	0.05	2
783	97	GLY	HA3	H	4.209	0.05	2
784	97	GLY	C	C	177.606	0.1	1
785	97	GLY	CA	C	46.630	0.1	1
786	97	GLY	N	N	112.164	0.05	1
787	98	PHE	H	H	9.599	0.02	1
788	98	PHE	HA	H	4.397	0.05	1
789	98	PHE	HB2	H	3.328	0.05	2
790	98	PHE	HB3	H	3.428	0.05	2
791	98	PHE	C	C	176.249	0.1	1
792	98	PHE	CA	C	61.648	0.1	1
793	98	PHE	CB	C	36.434	0.3	1
794	98	PHE	N	N	126.161	0.05	1
795	99	LEU	H	H	7.071	0.02	1
796	99	LEU	HA	H	4.071	0.05	1
797	99	LEU	HB2	H	1.398	0.05	2
798	99	LEU	HB3	H	1.699	0.05	2
799	99	LEU	HD2	H	0.773	0.05	2
800	99	LEU	C	C	178.874	0.1	1
801	99	LEU	CA	C	56.413	0.1	1
802	99	LEU	CB	C	43.234	0.3	1
803	99	LEU	CD2	C	23.468	0.1	2
804	99	LEU	N	N	121.449	0.05	1
805	100	SER	H	H	7.885	0.02	1
806	100	SER	HA	H	4.130	0.05	1
807	100	SER	HB2	H	4.022	0.05	2
808	100	SER	HB3	H	4.130	0.05	2
809	100	SER	C	C	174.439	0.1	1
810	100	SER	CA	C	59.486	0.1	1
811	100	SER	CB	C	62.347	0.3	1
812	100	SER	N	N	113.382	0.05	1
813	101	TRP	H	H	8.173	0.02	1
814	101	TRP	HA	H	4.042	0.05	1
815	101	TRP	HB2	H	3.495	0.05	2
816	101	TRP	HB3	H	3.687	0.05	2
817	101	TRP	HE1	H	8.953	0.05	1
818	101	TRP	C	C	176.430	0.1	1
819	101	TRP	CA	C	59.988	0.1	1
820	101	TRP	CB	C	27.775	0.3	1
821	101	TRP	N	N	122.914	0.05	1
822	101	TRP	NE1	N	127.740	0.05	1
823	102	VAL	H	H	5.549	0.02	1
824	102	VAL	HA	H	3.072	0.05	1
825	102	VAL	HB	H	1.862	0.05	1
826	102	VAL	HG1	H	-0.636	0.05	2
827	102	VAL	HG2	H	0.587	0.05	2
828	102	VAL	C	C	175.063	0.1	1
829	102	VAL	CA	C	63.524	0.1	1
830	102	VAL	CB	C	30.289	0.3	1
831	102	VAL	CG1	C	19.619	0.1	2
832	102	VAL	CG2	C	16.476	0.1	2
833	102	VAL	N	N	110.547	0.05	1
834	103	VAL	H	H	7.512	0.02	1
835	103	VAL	HA	H	4.712	0.05	1
836	103	VAL	HB	H	2.574	0.05	1
837	103	VAL	HG2	H	0.866	0.05	2

838	103	VAL	C	C	175.871	0.1	1
839	103	VAL	CA	C	60.527	0.1	1
840	103	VAL	CB	C	31.922	0.3	1
841	103	VAL	CG2	C	18.870	0.1	2
842	103	VAL	N	N	111.830	0.05	1
843	104	VAL	H	H	7.467	0.02	1
844	104	VAL	HA	H	4.057	0.05	1
845	104	VAL	HB	H	1.957	0.05	1
846	104	VAL	HG1	H	0.461	0.05	2
847	104	VAL	HG2	H	0.758	0.05	2
848	104	VAL	C	C	175.254	0.1	1
849	104	VAL	CA	C	63.259	0.1	1
850	104	VAL	CB	C	32.545	0.3	1
851	104	VAL	CG1	C	20.543	0.1	2
852	104	VAL	CG2	C	21.910	0.1	2
853	104	VAL	N	N	125.832	0.05	1
854	105	GLU	H	H	8.293	0.02	1
855	105	GLU	CA	C	53.206	0.1	1
856	105	GLU	N	N	126.035	0.05	1
857	106	PRO	HA	H	4.588	0.05	1
858	106	PRO	HB2	H	2.104	0.05	2
859	106	PRO	HB3	H	2.486	0.05	2
860	106	PRO	C	C	175.702	0.1	1
861	106	PRO	CA	C	64.175	0.1	1
862	106	PRO	CB	C	32.620	0.3	1
863	107	THR	H	H	7.267	0.02	1
864	107	THR	HA	H	5.028	0.05	1
865	107	THR	HB	H	4.514	0.05	1
866	107	THR	HG2	H	1.163	0.05	1
867	107	THR	CA	C	58.659	0.1	1
868	107	THR	CB	C	67.858	0.1	1
869	107	THR	CG2	C	21.344	0.1	1
870	107	THR	N	N	107.264	0.05	1
871	108	PRO	C	C	177.508	0.1	1
872	108	PRO	CA	C	65.925	0.1	1
873	108	PRO	CB	C	31.890	0.3	1
874	109	GLY	H	H	9.056	0.02	1
875	109	GLY	HA2	H	3.884	0.05	2
876	109	GLY	HA3	H	3.691	0.05	2
877	109	GLY	CA	C	48.547	0.1	1
878	109	GLY	N	N	109.064	0.05	1
879	110	PRO	C	C	174.425	0.1	1
880	110	PRO	CA	C	64.937	0.1	1
881	111	HIS	H	H	7.503	0.02	1
882	111	HIS	C	C	177.050	0.1	1
883	111	HIS	CA	C	60.755	0.1	1
884	111	HIS	CB	C	30.966	0.3	1
885	111	HIS	N	N	117.987	0.05	1
886	112	VAL	H	H	7.726	0.02	1
887	112	VAL	HA	H	3.389	0.05	1
888	112	VAL	HB	H	2.267	0.05	1
889	112	VAL	HG1	H	0.936	0.05	2
890	112	VAL	HG2	H	1.351	0.05	2
891	112	VAL	C	C	177.1	0.1	1
892	112	VAL	CA	C	68.371	0.1	1
893	112	VAL	CB	C	31.073	0.3	1
894	112	VAL	CG1	C	22.595	0.1	2
895	112	VAL	CG2	C	25.054	0.1	2
896	112	VAL	N	N	120.181	0.05	1
897	113	ALA	H	H	8.038	0.02	1
898	113	ALA	HA	H	3.998	0.05	1
899	113	ALA	HB	H	1.470	0.05	1
900	113	ALA	C	C	180.1	0.1	1
901	113	ALA	CA	C	55.671	0.1	1
902	113	ALA	CB	C	17.816	0.3	1
903	113	ALA	N	N	123.629	0.05	1
904	114	GLU	H	H	8.312	0.02	1
905	114	GLU	HA	H	4.081	0.05	1
906	114	GLU	HB2	H	2.245	0.05	2
907	114	GLU	HB3	H	2.398	0.05	2
908	114	GLU	C	C	179.421	0.1	1
909	114	GLU	CA	C	59.267	0.1	1
910	114	GLU	CB	C	29.011	0.3	1
911	114	GLU	N	N	122.183	0.05	1
912	115	MET	H	H	7.811	0.02	1
913	115	MET	HA	H	4.598	0.05	1
914	115	MET	HB3	H	2.399	0.05	2
915	115	MET	C	C	178.461	0.1	1
916	115	MET	CA	C	56.863	0.1	1
917	115	MET	CB	C	32.889	0.3	1
918	115	MET	N	N	121.498	0.05	1
919	116	ARG	H	H	8.967	0.02	1
920	116	ARG	HA	H	3.893	0.05	1
921	116	ARG	C	C	177.857	0.1	1
922	116	ARG	CA	C	59.662	0.1	1
923	116	ARG	CB	C	28.753	0.3	1
924	116	ARG	N	N	122.638	0.05	1
925	117	GLY	H	H	8.043	0.02	1
926	117	GLY	HA2	H	4.049	0.05	2
927	117	GLY	C	C	176.981	0.1	1
928	117	GLY	CA	C	47.410	0.1	1
929	117	GLY	N	N	108.279	0.05	1
930	118	SER	H	H	7.828	0.02	1
931	118	SER	HA	H	4.499	0.05	1
932	118	SER	HB2	H	4.299	0.05	2
933	118	SER	HB3	H	4.090	0.05	2
934	118	SER	C	C	176.975	0.1	1
935	118	SER	CA	C	62.226	0.1	1
936	118	SER	CB	C	63.464	0.3	1
937	118	SER	N	N	119.707	0.05	1
938	119	ALA	H	H	8.059	0.02	1
939	119	ALA	HA	H	4.677	0.05	2
940	119	ALA	HB	H	1.847	0.05	1
941	119	ALA	C	C	180.071	0.1	1
942	119	ALA	CA	C	55.562	0.1	1
943	119	ALA	CB	C	18.214	0.3	1
944	119	ALA	N	N	123.232	0.05	1
945	120	GLU	H	H	8.740	0.02	1
946	120	GLU	HA	H	5.453	0.05	1
947	120	GLU	HB2	H	2.028	0.05	2
948	120	GLU	HG2	H	2.582	0.05	2
949	120	GLU	HG3	H	2.226	0.05	2
950	120	GLU	C	C	178.166	0.1	1
951	120	GLU	CA	C	57.530	0.1	1
952	120	GLU	CB	C	29.451	0.3	1
953	120	GLU	CG	C	38.020	0.1	1
954	120	GLU	N	N	119.193	0.05	1
955	121	PHE	H	H	7.498	0.02	1
956	121	PHE	HA	H	4.126	0.05	1
957	121	PHE	HB2	H	3.249	0.05	2
958	121	PHE	HB3	H	2.936	0.05	2
959	121	PHE	HE1	H	6.485	0.05	3
960	121	PHE	C	C	177.255	0.1	1
961	121	PHE	CA	C	61.791	0.1	1
962	121	PHE	CB	C	39.356	0.3	1
963	121	PHE	N	N	120.925	0.05	1
964	122	TYR	H	H	7.413	0.02	1
965	122	TYR	HA	H	4.188	0.05	1
966	122	TYR	HB2	H	3.495	0.05	2
967	122	TYR	HB3	H	2.948	0.05	2
968	122	TYR	C	C	178.898	0.1	1
969	122	TYR	CA	C	61.970	0.1	1
970	122	TYR	CB	C	39.915	0.3	1
971	122	TYR	N	N	114.936	0.05	1

972	123	THR	H	H	9.730	0.02	1
973	123	THR	HA	H	4.210	0.05	1
974	123	THR	HB	H	4.732	0.05	1
975	123	THR	HG2	H	1.470	0.05	1
976	123	THR	C	C	177.902	0.1	1
977	123	THR	CA	C	66.916	0.1	1
978	123	THR	CB	C	66.480	0.1	1
979	123	THR	CG2	C	22.595	0.1	1
980	123	THR	N	N	115.420	0.05	1
981	124	ASN	H	H	9.013	0.02	1
982	124	ASN	HA	H	4.644	0.05	1
983	124	ASN	HB2	H	2.711	0.05	2
984	124	ASN	HB3	H	2.955	0.05	2
985	124	ASN	HD21	H	7.161	0.05	2
986	124	ASN	HD22	H	7.476	0.05	2
987	124	ASN	C	C	178.456	0.1	1
988	124	ASN	CA	C	56.135	0.1	1
989	124	ASN	CB	C	37.401	0.3	1
990	124	ASN	N	N	121.920	0.05	1
991	124	ASN	ND2	N	113.801	0.05	1
992	125	ARG	H	H	6.904	0.02	1
993	125	ARG	HA	H	4.020	0.05	1
994	125	ARG	HB2	H	1.961	0.05	2
995	125	ARG	HB3	H	2.149	0.05	2
996	125	ARG	HG2	H	1.537	0.05	2
997	125	ARG	HG3	H	1.753	0.05	2
998	125	ARG	C	C	178.409	0.1	1
999	125	ARG	CA	C	59.574	0.1	1
1000	125	ARG	CB	C	30.3	0.3	1
1001	125	ARG	N	N	120.64	0.05	1
1002	126	ILE	H	H	7.644	0.02	1
1003	126	ILE	HA	H	4.282	0.05	1
1004	126	ILE	HB	H	2.244	0.05	1
1005	126	ILE	HG2	H	1.093	0.05	1
1006	126	ILE	HD1	H	0.828	0.05	1
1007	126	ILE	C	C	178.20	0.1	1
1008	126	ILE	CA	C	65.083	0.1	1
1009	126	ILE	CB	C	38.679	0.3	1
1010	126	ILE	CG2	C	19.139	0.1	2
1011	126	ILE	N	N	123.198	0.05	1
1012	127	LEU	H	H	8.490	0.02	1
1013	127	LEU	HA	H	3.777	0.05	1
1014	127	LEU	HB2	H	2.291	0.05	2
1015	127	LEU	C	C	178.717	0.1	1
1016	127	LEU	CA	C	58.124	0.1	1
1017	127	LEU	CB	C	41.666	0.3	1
1018	127	LEU	N	N	119.405	0.05	1
1019	128	LYS	H	H	7.712	0.02	1
1020	128	LYS	HA	H	4.010	0.05	1
1021	128	LYS	HB2	H	1.944	0.05	2
1022	128	LYS	HG2	H	1.539	0.05	2
1023	128	LYS	HE2	H	3.013	0.05	2
1024	128	LYS	C	C	177.910	0.1	1
1025	128	LYS	CA	C	59.165	0.1	1
1026	128	LYS	CB	C	32.223	0.3	1
1027	128	LYS	N	N	119.161	0.05	1
1028	129	GLU	H	H	7.533	0.02	1
1029	129	GLU	HA	H	3.892	0.05	1
1030	129	GLU	HB2	H	1.515	0.05	2
1031	129	GLU	HB3	H	1.773	0.05	2
1032	129	GLU	C	C	177.864	0.1	1
1033	129	GLU	CA	C	58.673	0.1	1
1034	129	GLU	CB	C	30.332	0.3	1
1035	129	GLU	N	N	118.213	0.05	1
1036	130	PHE	H	H	8.368	0.02	1
1037	130	PHE	HA	H	4.894	0.05	1
1038	130	PHE	HB2	H	2.847	0.05	2
1039	130	PHE	HB3	H	3.621	0.05	2
1040	130	PHE	C	C	176.985	0.1	1
1041	130	PHE	CA	C	59.674	0.1	1
1042	130	PHE	CB	C	40.645	0.3	1
1043	130	PHE	N	N	112.505	0.05	1
1044	131	LYS	H	H	8.674	0.02	1
1045	131	LYS	HA	H	4.111	0.05	1
1046	131	LYS	HB2	H	1.548	0.05	2
1047	131	LYS	HB3	H	2.031	0.05	2
1048	131	LYS	C	C	177.421	0.1	1
1049	131	LYS	CA	C	59.684	0.1	1
1050	131	LYS	CB	C	31.406	0.3	1
1051	131	LYS	N	N	126.084	0.05	1
1052	132	GLY	H	H	8.939	0.02	1
1053	132	GLY	HA2	H	4.279	0.05	2
1054	132	GLY	HA3	H	3.837	0.05	2
1055	132	GLY	C	C	174.229	0.1	1
1056	132	GLY	CA	C	45.743	0.1	1
1057	132	GLY	N	N	116.744	0.05	1
1058	133	VAL	H	H	8.179	0.02	1
1059	133	VAL	HA	H	4.377	0.05	1
1060	133	VAL	HB	H	2.122	0.05	1
1061	133	VAL	HG1	H	1.000	0.05	2
1062	133	VAL	HG2	H	1.189	0.05	2
1063	133	VAL	C	C	175.801	0.1	1
1064	133	VAL	CA	C	63.638	0.1	1
1065	133	VAL	CB	C	35.220	0.3	1
1066	133	VAL	N	N	119.830	0.05	1
1067	134	ASN	H	H	8.543	0.02	1
1068	134	ASN	HA	H	5.065	0.05	1
1069	134	ASN	HB3	H	2.932	0.05	2
1070	134	ASN	C	C	174.745	0.1	1
1071	134	ASN	CA	C	52.457	0.1	1
1072	134	ASN	CB	C	40.388	0.3	1
1073	134	ASN	N	N	118.614	0.05	1
1074	135	GLN	H	H	9.018	0.02	1
1075	135	GLN	HA	H	3.930	0.05	1
1076	135	GLN	HB2	H	2.291	0.05	2
1077	135	GLN	HB3	H	2.747	0.05	2
1078	135	GLN	C	C	176.762	0.1	1
1079	135	GLN	CA	C	57.330	0.1	1
1080	135	GLN	CB	C	28.463	0.3	1
1081	135	GLN	N	N	127.879	0.05	1
1082	136	ASP	H	H	7.984	0.02	1
1083	136	ASP	HA	H	4.383	0.05	1
1084	136	ASP	HB2	H	2.425	0.05	2
1085	136	ASP	HB3	H	2.945	0.05	2
1086	136	ASP	C	C	178.318	0.1	1
1087	136	ASP	CA	C	57.848	0.1	1
1088	136	ASP	CB	C	39.593	0.3	1
1089	136	ASP	N	N	119.727	0.05	1
1090	137	GLN	H	H	7.201	0.02	1
1091	137	GLN	HA	H	4.367	0.05	1
1092	137	GLN	HB2	H	1.844	0.05	2
1093	137	GLN	HB3	H	1.674	0.05	2
1094	137	GLN	HE21	H	6.073	0.05	2
1095	137	GLN	C	C	177.606	0.1	1
1096	137	GLN	CA	C	59.004	0.1	1
1097	137	GLN	CB	C	28.688	0.3	1
1098	137	GLN	N	N	117.279	0.05	1
1099	137	GLN	NE2	N	106.538	0.05	1
1100	138	VAL	H	H	6.907	0.02	1
1101	138	VAL	HA	H	3.565	0.05	1
1102	138	VAL	HB	H	2.301	0.05	1
1103	138	VAL	HG1	H	1.096	0.05	2
1104	138	VAL	HG2	H	1.163	0.05	2
1105	138	VAL	C	C	178.261	0.1	1

1106	138	VAL	CA	C	66.636	0.1	1
1107	138	VAL	CB	C	31.481	0.3	1
1108	138	VAL	N	N	119.102	0.05	1
1109	139	ASP	H	H	8.888	0.02	1
1110	139	ASP	HA	H	4.490	0.05	1
1111	139	ASP	HB2	H	2.755	0.05	2
1112	139	ASP	HB3	H	2.916	0.05	2
1113	139	ASP	C	C	177.704	0.1	1
1114	139	ASP	CA	C	57.321	0.1	1
1115	139	ASP	CB	C	40.345	0.3	1
1116	139	ASP	N	N	122.560	0.05	1
1117	140	TRP	H	H	8.574	0.02	1
1118	140	TRP	HA	H	4.093	0.05	1
1119	140	TRP	HB2	H	3.906	0.05	2
1120	140	TRP	HB3	H	3.417	0.05	2
1121	140	TRP	HE1	H	9.107	0.05	1
1122	140	TRP	C	C	177.147	0.1	1
1123	140	TRP	CA	C	63.988	0.1	1
1124	140	TRP	CB	C	29.344	0.3	1
1125	140	TRP	N	N	125.972	0.05	1
1126	140	TRP	NE1	N	132.523	0.05	1
1127	141	VAL	H	H	8.220	0.02	1
1128	141	VAL	HA	H	3.571	0.05	1
1129	141	VAL	HB	H	2.673	0.05	1
1130	141	VAL	HG1	H	1.191	0.05	2
1131	141	VAL	HG2	H	1.634	0.05	2
1132	141	VAL	C	C	177.209	0.1	1
1133	141	VAL	CA	C	67.780	0.1	1
1134	141	VAL	CB	C	32.072	0.3	1
1135	141	VAL	CG1	C	22.595	0.1	2
1136	141	VAL	CG2	C	24.293	0.1	2
1137	141	VAL	N	N	117.702	0.05	1
1138	142	SER	H	H	8.313	0.02	1
1139	142	SER	HA	H	4.098	0.05	1
1140	142	SER	HB3	H	4.001	0.05	2
1141	142	SER	C	C	177.554	0.1	1
1142	142	SER	CA	C	61.740	0.1	1
1143	142	SER	CB	C	63.421	0.3	1
1144	142	SER	N	N	113.762	0.05	1
1145	143	ASN	H	H	8.157	0.02	1
1146	143	ASN	HA	H	4.628	0.05	1
1147	143	ASN	HB2	H	2.732	0.05	2
1148	143	ASN	HB3	H	3.145	0.05	2
1149	143	ASN	C	C	176.361	0.1	1
1150	143	ASN	CA	C	56.738	0.1	1
1151	143	ASN	CB	C	39.808	0.3	1
1152	143	ASN	N	N	118.388	0.05	1
1153	144	TYR	H	H	7.287	0.02	1
1154	144	TYR	HA	H	3.413	0.05	1
1155	144	TYR	HB2	H	2.202	0.05	2
1156	144	TYR	HB3	H	2.101	0.05	2
1157	144	TYR	HE1	H	6.132	0.05	3
1158	144	TYR	C	C	176.792	0.1	1
1159	144	TYR	CA	C	61.613	0.1	1
1160	144	TYR	CB	C	38.024	0.3	1
1161	144	TYR	N	N	119.866	0.05	1
1162	145	VAL	H	H	8.417	0.02	1
1163	145	VAL	HA	H	3.084	0.05	1
1164	145	VAL	HB	H	2.034	0.05	1
1165	145	VAL	HG1	H	0.884	0.05	2
1166	145	VAL	HG2	H	1.092	0.05	2
1167	145	VAL	C	C	177.862	0.1	1
1168	145	VAL	CA	C	67.454	0.1	1
1169	145	VAL	CB	C	31.234	0.3	1
1170	145	VAL	N	N	117.103	0.05	1
1171	146	ASN	H	H	8.357	0.02	1
1172	146	ASN	HA	H	4.164	0.05	1
1173	146	ASN	HB2	H	2.850	0.05	2
1174	146	ASN	HB3	H	2.251	0.05	2
1175	146	ASN	HD21	H	7.692	0.05	2
1176	146	ASN	C	C	177.290	0.1	1
1177	146	ASN	CA	C	55.689	0.1	1
1178	146	ASN	CB	C	37.122	0.3	1
1179	146	ASN	N	N	118.276	0.05	1
1180	146	ASN	ND2	N	112.775	0.05	1
1181	147	PHE	H	H	7.429	0.02	1
1182	147	PHE	HA	H	3.485	0.05	1
1183	147	PHE	HB2	H	2.779	0.05	2
1184	147	PHE	HB3	H	2.889	0.05	2
1185	147	PHE	C	C	175.795	0.1	1
1186	147	PHE	CA	C	61.557	0.1	1
1187	147	PHE	CB	C	37.423	0.3	1
1188	147	PHE	N	N	121.987	0.05	1
1189	148	LEU	H	H	7.235	0.02	1
1190	148	LEU	HA	H	3.201	0.05	1
1191	148	LEU	HB2	H	0.761	0.05	2
1192	148	LEU	HB3	H	1.516	0.05	2
1193	148	LEU	HG	H	0.931	0.05	1
1194	148	LEU	HD1	H	-0.266	0.05	2
1195	148	LEU	HD2	H	0.295	0.05	2
1196	148	LEU	C	C	177.3	0.1	1
1197	148	LEU	CA	C	57.542	0.1	1
1198	148	LEU	CB	C	41.784	0.3	1
1199	148	LEU	CD1	C	20.827	0.1	2
1200	148	LEU	CD2	C	25.054	0.1	2
1201	148	LEU	N	N	118.175	0.05	1
1202	149	LYS	H	H	8.308	0.02	1
1203	149	LYS	HA	H	4.425	0.05	1
1204	149	LYS	C	C	178.880	0.1	1
1205	149	LYS	CA	C	60.441	0.1	1
1206	149	LYS	CB	C	32.438	0.3	1
1207	149	LYS	N	N	122.049	0.05	1
1208	150	ASP	H	H	8.286	0.02	1
1209	150	ASP	HA	H	4.556	0.05	1
1210	150	ASP	HB2	H	2.624	0.05	2
1211	150	ASP	HB3	H	1.989	0.05	2
1212	150	ASP	C	C	178.835	0.1	1
1213	150	ASP	CA	C	56.584	0.1	1
1214	150	ASP	CB	C	39.249	0.3	1
1215	150	ASP	N	N	122.986	0.05	1
1216	151	LEU	H	H	8.852	0.02	1
1217	151	LEU	HA	H	3.938	0.05	1
1218	151	LEU	HG	H	1.523	0.05	1
1219	151	LEU	HD1	H	0.623	0.05	2
1220	151	LEU	HD2	H	0.977	0.05	2
1221	151	LEU	C	C	177.6	0.1	1
1222	151	LEU	CA	C	57.441	0.1	1
1223	151	LEU	CB	C	40.205	0.3	1
1224	151	LEU	CD1	C	25.753	0.1	2
1225	151	LEU	N	N	126.639	0.05	1
1226	152	GLU	H	H	7.51	0.02	1
1227	152	GLU	HA	H	3.610	0.05	1
1228	152	GLU	HB2	H	1.902	0.05	2
1229	152	GLU	HB3	H	2.105	0.05	2
1230	152	GLU	HG2	H	2.309	0.05	2
1231	152	GLU	C	C	177.425	0.1	1
1232	152	GLU	CA	C	60.348	0.1	1
1233	152	GLU	CB	C	29.440	0.3	1
1234	152	GLU	N	N	121.179	0.05	1
1235	153	LYS	H	H	7.238	0.02	1
1236	153	LYS	HA	H	3.896	0.05	1
1237	153	LYS	HB2	H	1.656	0,05	2
1238	153	LYS	HB3	H	1.988	0.05	2
1239	153	LYS	HG2	H	1.429	0.05	2

1240	153	LYS	HD2	H	1.669	0.05	2
1241	153	LYS	C	C	177.939	0.1	1
1242	153	LYS	CA	C	60.212	0.1	1
1243	153	LYS	CB	C	32.470	0.3	1
1244	153	LYS	N	N	119.686	0.05	1
1245	154	TYR	H	H	8.322	0.02	1
1246	154	TYR	HA	H	4.420	0.05	1
1247	154	TYR	HB3	H	3.432	0.05	2
1248	154	TYR	C	C	176.936	0.1	1
1249	154	TYR	CA	C	61.271	0.1	1
1250	154	TYR	CB	C	38.357	0.3	1
1251	154	TYR	N	N	124.041	0.05	1
1252	155	ILE	H	H	8.475	0.02	1
1253	155	ILE	HA	H	2.984	0.05	1
1254	155	ILE	HB	H	1.561	0.05	1
1255	155	ILE	HG12	H	-0.804	0.05	2
1256	155	ILE	HG13	H	1.253	0.05	2
1257	155	ILE	HG2	H	0.577	0.05	1
1258	155	ILE	HD1	H	-0.108	0.05	1
1259	155	ILE	C	C	177.556	0.1	1
1260	155	ILE	CA	C	65.533	0.1	1
1261	155	ILE	CB	C	38.615	0.3	1
1262	155	ILE	CG1	C	28.933	0.1	2
1263	155	ILE	CG2	C	17.846	0.1	2
1264	155	ILE	CD1	C	16.767	0.1	1
1265	155	ILE	N	N	123.426	0.05	1
1266	156	LYS	H	H	7.979	0.02	1
1267	156	LYS	HA	H	3.843	0.05	1
1268	156	LYS	HB2	H	1.920	0.05	2
1269	156	LYS	HG2	H	1.471	0.05	2
1270	156	LYS	HD2	H	1.662	0.05	2
1271	156	LYS	C	C	177.964	0.1	1
1272	156	LYS	CA	C	60.050	0.1	1
1273	156	LYS	CB	C	31.782	0.3	1
1274	156	LYS	N	N	118.659	0.05	1
1275	157	GLN	H	H	7.666	0.02	1
1276	157	GLN	HA	H	3.971	0.05	1
1277	157	GLN	C	C	176.987	0.1	1
1278	157	GLN	CA	C	58.162	0.1	1
1279	157	GLN	CB	C	28.946	0.3	1
1280	157	GLN	N	N	119.245	0.05	1
1281	158	TYR	H	H	7.940	0.02	1
1282	158	TYR	HA	H	4.233	0.05	1
1283	158	TYR	HB2	H	2.318	0.05	2
1284	158	TYR	HB3	H	1.716	0.05	2
1285	158	TYR	C	C	175.640	0.1	1
1286	158	TYR	CA	C	60.130	0.1	1
1287	158	TYR	CB	C	39.990	0.3	1
1288	158	TYR	N	N	115.134	0.05	1
1289	159	HIS	H	H	8.534	0.02	1
1290	159	HIS	HA	H	4.985	0.05	1
1291	159	HIS	HB2	H	3.507	0.05	2
1292	159	HIS	HB3	H	2.570	0.05	2
1293	159	HIS	C	C	176.615	0.1	1
1294	159	HIS	CA	C	54.178	0.1	1
1295	159	HIS	CB	C	32.943	0.3	1
1296	159	HIS	N	N	121.458	0.05	1
1297	160	THR	H	H	7.284	0.02	1
1298	160	THR	HA	H	3.694	0.05	1
1299	160	THR	HB	H	4.036	0.05	1
1300	160	THR	HG2	H	1.446	0.05	1
1301	160	THR	C	C	176.155	0.1	1
1302	160	THR	CA	C	68.492	0.1	1
1303	160	THR	CB	C	70.114	0.3	1
1304	160	THR	N	N	113.903	0.05	1
1305	161	THR	H	H	8.161	0.02	1
1306	161	THR	HA	H	5.278	0.05	1
1307	161	THR	HB	H	4.863	0.05	1
1308	161	THR	HG2	H	1.330	0.05	1
1309	161	THR	C	C	174.601	0.1	1
1310	161	THR	CA	C	60.122	0.1	1
1311	161	THR	CB	C	68.524	0.3	1
1312	161	THR	CG2	C	21.904	0.1	1
1313	161	THR	N	N	107.226	0.05	1
1314	162	GLY	H	H	7.674	0.02	1
1315	162	GLY	HA2	H	4.703	0.05	2
1316	162	GLY	HA3	H	3.945	0.05	2
1317	162	GLY	C	C	173.474	0.1	1
1318	162	GLY	CA	C	43.994	0.1	1
1319	162	GLY	N	N	111.699	0.05	1
1320	163	LEU	H	H	8.894	0.02	1
1321	163	LEU	HA	H	4.245	0.05	1
1322	163	LEU	HB2	H	1.898	0.05	2
1323	163	LEU	HB3	H	1.117	0.05	2
1324	163	LEU	HG	H	0.671	0.05	1
1325	163	LEU	HD1	H	0.347	0.05	2
1326	163	LEU	HD2	H	0.322	0.05	2
1327	163	LEU	C	C	176.467	0.1	1
1328	163	LEU	CA	C	56.194	0.1	1
1329	163	LEU	CB	C	42.354	0.3	1
1330	163	LEU	CD1	C	24.520	0.3	1
1331	163	LEU	N	N	124.062	0.05	1
1332	164	THR	H	H	10.375	0.02	1
1333	164	THR	HA	H	4.164	0.05	1
1334	164	THR	HB	H	4.298	0.05	1
1335	164	THR	HG2	H	1.400	0.05	1
1336	164	THR	C	C	174.340	0.1	1
1337	164	THR	CA	C	64.671	0.1	1
1338	164	THR	CB	C	69.212	0.3	1
1339	164	THR	N	N	129.322	0.05	1
1340	165	TRP	H	H	8.158	0.02	1
1341	165	TRP	HA	H	5.125	0.05	1
1342	165	TRP	HB2	H	3.719	0.05	2
1343	165	TRP	HB3	H	2.938	0.05	2
1344	165	TRP	HE1	H	10.360	0.05	1
1345	165	TRP	C	C	175.853	0.1	1
1346	165	TRP	CA	C	56.780	0.1	1
1347	165	TRP	CB	C	29.677	0.3	1
1348	165	TRP	N	N	130.840	0.05	1
1349	165	TRP	NE1	N	130.786	0.05	1
1350	166	ASN	H	H	8.777	0.02	1
1351	166	ASN	HA	H	5.245	0.05	1
1352	166	ASN	HB2	H	2.874	0.05	2
1353	166	ASN	HB3	H	2.710	0.05	2
1354	166	ASN	CA	C	49.653	0.1	1
1355	166	ASN	CB	C	39.403	0.1	1
1356	166	ASN	N	N	121.099	0.05	1
1357	167	PRO	HA	H	5.242	0.05	1
1358	167	PRO	HB2	H	2.357	0.05	2
1359	167	PRO	HB3	H	2.570	0.05	2
1360	167	PRO	HG3	H	2.198	0.05	2
1361	167	PRO	C	C	177.480	0.1	1
1362	167	PRO	CA	C	64.399	0.1	1
1363	167	PRO	CB	C	32.298	0.3	1
1364	168	LYS	H	H	7.716	0.02	1
1365	168	LYS	HA	H	4.555	0.05	1
1366	168	LYS	HB2	H	2.149	0.05	2
1367	168	LYS	HB3	H	1.753	0.05	2
1368	168	LYS	HG2	H	1.459	0.05	2
1369	168	LYS	C	C	176.938	0.1	1
1370	168	LYS	CA	C	55.176	0.1	1
1371	168	LYS	CB	C	31.847	0.3	1
1372	168	LYS	N	N	117.196	0.05	1
1373	169	GLY	H	H	7.976	0.02	1

```
 1374   169   GLY  HA2   H    3.573     0.05  2
 1375   169   GLY  HA3   H    4.253     0.05  2
 1376   169   GLY  C     C  173.879     0.1   1
 1377   169   GLY  CA    C   44.400     0.1   1
 1378   169   GLY  N     N  109.825     0.05  1
 1379   170   GLY  H     H    7.984     0.02  1
 1380   170   GLY  HA2   H    3.887     0.05  2
 1381   170   GLY  HA3   H    4.514     0.05  2
 1382   170   GLY  C     C  173.963     0.1   1
 1383   170   GLY  CA    C   44.314     0.1   1
 1384   170   GLY  N     N  106.896     0.05  1
 1385   171   ASP  H     H    8.343     0.02  1
 1386   171   ASP  HA    H    4.806     0.05  1
 1387   171   ASP  HB2   H    2.739     0.05  2
 1388   171   ASP  HB3   H    2.569     0.05  2
 1389   171   ASP  C     C  177.287     0.1   1
 1390   171   ASP  CA    C   54.040     0.1   1
 1391   171   ASP  CB    C   41.902     0.3   1
 1392   171   ASP  N     N  120.195     0.05  1
 1393   172   ALA  H     H    9.461     0.02  1
 1394   172   ALA  HA    H    4.372     0.05  1
 1395   172   ALA  HB    H    1.468     0.05  1
 1396   172   ALA  C     C  177.346     0.1   1
 1397   172   ALA  CA    C   53.495     0.1   1
 1398   172   ALA  CB    C   18.869     0.3   1
 1399   172   ALA  N     N  130.180     0.05  1
 1400   173   LYS  H     H    8.273     0.02  1
 1401   173   LYS  HA    H    4.324     0.05  1
 1402   173   LYS  HB2   H    1.785     0.05  2
 1403   173   LYS  HB3   H    1.922     0.05  2
 1404   173   LYS  HG2   H    0.936     0.05  2
 1405   173   LYS  HG3   H    1.164     0.05  2
 1406   173   LYS  C     C  180.379     0.1   1
 1407   173   LYS  CA    C   56.451     0.1   1
 1408   173   LYS  CB    C   33.157     0.3   1
 1409   173   LYS  N     N  118.737     0.05  1
 1410   174   SER  H     H    8.173     0.02  1
 1411   174   SER  HA    H    4.469     0.05  1
 1412   174   SER  HB2   H    4.328     0.05  2
 1413   174   SER  HB3   H    3.952     0.05  2
 1414   174   SER  C     C  174.517     0.1   1
 1415   174   SER  CA    C   58.338     0.1   1
 1416   174   SER  CB    C   63.776     0.3   1
 1417   174   SER  N     N  117.133     0.05  1
 1418   175   ALA  H     H    8.072     0.02  1
 1419   175   ALA  HA    H    4.451     0.05  1
 1420   175   ALA  HB    H    1.485     0.05  1
 1421   175   ALA  C     C  176.862     0.1   1
 1422   175   ALA  CA    C   52.623     0.1   1
 1423   175   ALA  CB    C   19.385     0.3   1
 1424   175   ALA  N     N  127.342     0.05  1
 1425   176   THR  H     H    7.765     0.02  1
 1426   176   THR  HA    H    4.196     0.05  1
 1427   176   THR  HG2   H    1.212     0.05  1
 1428   176   THR  CA    C   63.099     0.1   1
 1429   176   THR  N     N  119.501     0.05  1

stop_

save_
```

Bibliography

Abragam A.: *Principles of nuclear mechanism*, Oxford University Press, (1961)

Anglister J., Grzesiek S., Ren H., Klee C. & Bax A.: Isotope-edited multidimensional NMR of calcineurin B in the presence of the non-deuterated detergent CHAPS, *J. Biomol. NMR.* **3** (1993) 121–126

Baumgartner R., Fernandez-Catalan C., Winoto A., Huber R., Engh R. & T.A. H.: Stucture of human cyclin-dependent kinase inhibitor P19^{INK4d} - comparison to known ankyrin-repeat-containing structures and implication for the dysfunktion of tumor suppressor p16^{INK4d}, *Structure* **6** (1998) 1279–1290

Baxter R.C.: Signalling pathways involved in antiproliferative effects of IGFBP-3: a review, *Mol. Pathol.* **54** (2001) 145–148

Baxter R.C., Bayne M.L. & Cascieri M.A.: Structural determinants for binary and ternary complex formation between insulin-like growth factor-I (IGF-I) and IGF binding protein-3, *J. Biol. Chem.* **267** (1992) 60–65

Bayne M.L., Applebaum J., Chicchi G.G., Miller R.E. & Cassieri M.A.: The roles of tyrosines 24, 31, and 60 in the high affinity binding of insulin-like growth factor-I to the type 1 insuline-like growth factor receptor, *J. Biol. Chem.* **265** (1990) 15648–15652

Berman H., Westbrook J., Feng Z., Gilliland G., Bhat T., Weissig H., Shindyalov I. & Bourne P.: The Protein Data Bank, *Nucleic Acids Research* **28** (2000) 235–242

Brocks B., Rode H., Klein M., Gerlach E., Dubel S., Little M., Pfizermaier K. & Moosmayer D.: A TNF receptor antagonistic scFv, which is not secreted in mammalian cells, is expressed as a soluble mono- and bivalent scFv derivative in insect cells, *Immunotechnology* **3** (1997) 173–184

Brüggert M.: *NMR-spektroskopische Untersuchungen zur Expression selektiv mit ^{15}N-Aminosäuren markierter Proteine in* Spodoptera frugiperda *(Sf9) Insektenzellen*, Doktorarbeit, Technische Universität München, Fakultät für Chemie, (2002)

Canet D.: *Nuclear Magnetic Resonance. Concepts and Methods*, John Wiley & Sons, (1996)

Canet D., Last A., Tito P., Sunde M., Spencer A., Archer D., Redfield C., Robinson C. & C.M. D.: Local cooperativity in the unfolding of an amyloidogenic variant of human lysozyme, *Nat. Struct. Biol.* **9** (2002) 308–315

Cascio M.: *Methods in Neuroscience*, Academic Press, (1995)

Cavanagh J., Fairbrother W., Palmer III A. & Skelton N.: *Protein NMR spectroscopy: Principles and Practice*, Academic Press, (1996)

Chen J., Marechal V. & Levine A.J.: Mapping of the p53 and MDM2 interaction domains, *Mol. Cell. Biol.* **13** (1993) 4107–4114

Chiou T.W., Hsieh Y.C. & Ho C.: High density culture of insect cells using rational medium design and feeding strategy, *Bioprocess Engeneering* **22** (2000) 483–491

Christendat D., Yee A., Dharamsi A., Kluger Y., Gerstein M., Arrowsmith C. & Edwards A.: Structural proteomics: prospects for high throughput sample preparation, *Prog. Biophys. Mol. Biol.* **73** (2000) 339–345

Cieslar C., Clore M. & Gronenborn J.: Computer-Aided Sequential Assignment of Protein ^{1}H NMR Spectra, *J. Magn. Reson.* **80** (1988) 119

Cieslar C., Holak T. & Oschkinat H.: A Program for the Evaluation of 3d Spectra Applied to the Sequential Assignment of BPTI Utilizing 3D TOCSY-NOESY, *J. Magn. Reson.* **87** (1990) 400–407

Clemmons D.R.: Use of mutagenesis to probe IGF-binding protein structure/function relationships, *Endocr. Rev.* **22** (2001) 800–817

Crameri A., Whitehorn E., Tate E. & Stemmer W.: Improved green fluorescent protein by molecular evolution using DNA shuffling, *Nat. Biotechnol.* **14** (1996) 315–319

Creemers A., Klaasen C., Bovee-Geurts P., Kelle R., Kragl U., Raap J., de Grip W. & de Groot H.: Solid State ^{15}NNMR Evidence for a Complex Schiff Base Counterion in the Visual G-Protein-Coupled Receptor Rhodopsin, *Biochemistry* **38** (1999) 7195–7199

Croasmun W. & Carlson R.: *Two dimensional NMR Spectroscopy. Applications for Chemists and Biochemists*, VCH Publisher, (1994)

Daskiewicz J.B., Comte G., Barron D., Di Petro A. & Thomasson F.: Organolithium mediated synthesis of prenylchalcones as potential inhibitors of chemoresistance, *Tetrahedron Lett.* **40** (1999) 7095–7098

DeLange A., Klaasen C., Wallace-Williams S., Bovee-Geurts P., Liu X.M., de Grip W. & Rothschild K.: Tyrosine Structural Changes Detected during the Photoactivation of Rhodopsin, *J. Biol. Chem.* **273** (1998) 23735–23739

Dempsey C.: Hydrogen exchange in peptides and proteins using NMR spectroscopy, *Progress in Nuclear Magnetic Resonance Spectroscopy* **39** (2001) 135–170

Derome A.: *Modern NMR Techniques for Chemistry Research*, Pergamon Press, (1987)

Devincenzo R., Scambia G., Panici P.B., Ranelletti F.O., Bonanno G., Ercoli A., Dellemonache F., Ferrari F., Piantelli M. & Mancuso S.: Effect of synthetic and naturally occuring chalcones on ovarian cancer cell growth - structure-activity relationships, *Anti-Cancer Drug Des.* **10** (1995) 481–490

Diercks T., Coles M. & Kessler H.: Applications of NMR in drug discovery, *Curr. Opin. Chem. Biol.* **5** (2001) 285–291

Doverskog M., Han L. & Häggström L.: Cystine/cysteine metabolism in cultured Sf9 cells: influence of cell physiology on biosynthesis, amino acid uptake and growth, *Cytotechnology* **26** (1998) 91–102

Doverskog M., Jacobsson U., Chapman B., Kuchel P. & Häggström L.: Determination of NADH-dependent glutamate synthase (GOGAT) in Spodoptera frugiperda (Sf9) insect cells by a selective 1H/15N NMR in vitro assay, *J. Biotechnol.* **79** (2000) 87–97

Drews M., T. P. & Vilu R.: The growth and nutrient utilization of the insect cell line Spodoptera frugiperda Sf9 in batch and continuous culture, *J. Biotechnol.* **40** (1995) 187–198

Drews M., Doverskog M., Öhman L., Chapman B., Jacobsson U., Kuchel P. & Häggström L.: Pathways of glutamine metabolism in Spodoptera frugiperda (Sf9) insect cells: evidence for the presence of the nitrogen assimilation system, and a metabolic switch by $^{1}H/^{15}N$ NMR, *J. Biotechnol.* **78** (2000) 23–37

Edwards A., Arrowsmith C., Christendat D., Dharamsi A., Friesen J., Greenblatt J. & Vedadi M.: Protein production: feeding the crystallographers and NMR spectroscopists., *Nat. Struct. Biol.* **struct. genomics suppl.** (2000) 970–972

Ernst R., Bodenhausen G. & Wokaun A.: *Principles of Magnetic Resonance in One and Two Dimensions*, Clarendon Press, (1997)

Evans J.: *Biomolecular NMR spectroscopy*, Oxford Univ. Press, (1995)

Ferrance J., Goel A. & Ataai M.: Utilization of glucose and amino acids in insect cell cultures: quantifying the metabolic flows within primary pathways and medium development, *Biotechnol. Bioeng.* **42** (1993) 697–707

Field J., Vojtek A., Ballester R., Bolger G., Colicelli J., Ferguson K., Gerst J., Kataoka T., Michaeli T., Powers S., Riggs M., Rodgers L., Wieland I., Wheland B. & Wigler M.: Cloning and characterization of CAP, the S. cerevisiae gene encoding the 70 kd adenylyl cyclase-associated protein, *Cell* **61** (1990) 319–327

Fielding L.: Determination of the association constants (Ka) from solution NMR data, *Tetrahedon* **56** (2000) 6151–6170

Fischer M., Majumdar A. & Zuiderweg E.: Protein NMR relaxation: Theory, applications and outlook, *Prog. Nucl. Magn. Res. Spec.* **33** (1998) 207–272

Freeman R.: *A Handbook of Nuclear Magnetic Resonance*, Longman, (1988)

Georgescu J.: *NMR-Untersuchungen und Strukturbestimmungen von Biomolekülen: Linker-polypeptid $L_C^{7,8}$, Interleukin-16, grün fluoreszierendes Protein (GFPuv)*, Doktorarbeit, Technische Universität München, Fakultät für Chemie, (2000)

Gerst J., Rodgers L., Riggs M. & Wigler M.: SNC1, a yeast homolog of the synaptic vesicle-associated membrane protein/synaptobrevin gene family: genetic interactions with the RAS and CAP genes, *Proc. Natl. Acad. Sci.* **89** (1992)

Gluckman P., Klempt N., Guan J., Mallard C., Sirimanne E., Dragunow M., Singh K., Williams C. & Nikolics K.: A role for IGF-1 in the rescue of CNS neurons following hypoxic-ischemic injury, *Biochem. Biophys. Res. Commun.* **182** (1992) 593–599

Goddard T. & Kneller D.: SPARKY 3, University of California, San Francisco

Gottwald U., Brokamp R., Karakesisoglou I., Schleicher M. & Noegel A.: Identification of a cyclase-associated protein (CAP) homologue in Dictyostelium dicoideum and characterization of Its Interaction with actin, *Mol. Biol. Cell* **7** (1996) 261–272

Hankinson S.E., Willet W.C., Colditz G.A., Hunter D.J., Michaud D.S., Deroo B., Rosner B., Speizer F.E. & Pollak M.: Circulating concentrations of insulin-like growth factor I and risk of breast cancer, *Lancet* **351** (1998) 1393–1396

Haupt Y., Maya R., Kazaz A. & Oren M.: MDM2 promotes the rapid degradation of p53, *Nature* **387** (1997) 296–299

Haupts U., Maiti S., Schwille P. & Webb W.: Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy, *Proc. Nat. Acad. Sci. USA* **95** (1998) 13573–13578

Holly J.: Insulin-like growth factor-I and new opportunities for cancer prevention, *Lancet* **351** (2000) 1373–1374

Hubberstey A. & Mottillio E.: Cyclase-associated proteins: CAPacity for linking signal transduction and actin polymerization, *FASEB* **16** (2002) 487–499

Hwa V. et al.: The IGF binding protein superfamily, in: *The IGF system. Molecular biology, physiology, and clinical applications* (Eds. Rosenfeld R.G. & Roberts C.T.), Humana Press: Totowa, 1999 S. 315–327

James T. & Oppenheimer N. (Eds.): *Nuclear Magnetic Resonance*, Academic Press, (1994)

James T., Dötsch V. & Schmitz U. (Eds.): *Nuclear Magnetic Resonance of Biological Macromolecules, Part A*, Academic Press, (2001)

Jones J.L. & Clemmons D.R.: Insulin-like growth factors and their binding proteins: biological actions, *Endocr. Rev.* **12** (1995) 10–21

Juven-Gershon T. & Oren M.: Mdm2: The ups and downs, *Mol. Med.* **5** (1999) 71–83

Kalus W., Baumgartner R., Renner C., Noegel A., Chan F., Winoto A. & Holak T.: NMR structural characterization of the CDK inhibitor P19^{INK4d}, *FEBS* **401** (1997) 127–132

Kalus W., Zweckstetter M., Renner C., Sanchez Y., Georgescu J., Grol M., Demuth D., Schumacher R., Dony C., Lang K. & Holak T.: Structure of the IGF-binding domain of the insulin-like growth factor-binding protein-5 (IGFBP-5): implications for IGF and IGF-I receptor interactions, *EMBO J.* **17** (1998) 6558–6572

Kay L.: Protein dynamics from NMR, *Nature Struct. Biol.* **NMR supplement** (1998) 513–517

Khandwala H.M., McCutcheon I.E., Flyvbjerg A. & Friend K.E.: The effects of insulin-like growth factors on tumorigenesis and neoplastic growth, *Endocr. Rev.* **21** (2000) 215–244

Kim J., Kim E. & Park T.: Fed-batch culture of insect cells with exponential feeding of amino acid and yeastolate solution, *Bioprocess Engineering* **23** (2000) 367–370

Ksiazek D.: *Structural investigations on medically relevant proteins*, Doktorarbeit, Technische Universität München, Fakultät für Chemie, (2002)

Kubbutat M.H.G., Jones S.N. & Vousden K.H.: Regulation of p53 stability by MDM2, *Nature* **387** (1997) 299–303

Kumar C., Diao R., Yin Y., Samatar A., Madison V. & Xiao L.: Expression, purification, characterization and homology modeling of active Akt/PKB, a key enzyme involved in cell survival signaling, *Biochim. Biophys. Acta-Gen. Subj.* **1526** (2001) 257–268

Kussie P.H., Gorina S., Marechal V., Elenbaas B., Moreau J., Levine A.J. & Pavletich N.P.: Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain, *Science* **274** (1996) 948–953.

Labas Y., Gurskaya N., Yanushevich Y., Fradkov A., Lukyanov K., Lukyanov S. & Matz M.: Diversity and evolution of the green fluorescent protein family, *Proc. Nat. Acad. Sci. USA* **99** (2002) 4256–4261

Lane D.P. & Hall P.A.: MDM2 - arbiter of p53's destruction, *Trends Biochem. Sci.* **22** (1997) 372–374

Lawrie A., King L. & Ogden J.: High level synthesis and secretion of human urokinase using a late gene promoter of the Autographa californica nuclear polyhedrosis virus, *J. Biotechnol.* **39** (1995) 1–8

Liu X.J., Xie Q., Zhu Y.F., Chen C. & Ling N.: Identification of a nonpeptide ligand that releases bioactive insulin-like growth factor-I from its binding protein complex, *J. Biol. Chem.* **276** (2001) 32419–32422

Loddick S.A., Liu X.J., Lu Z.X., Liu C.L., Behan D.P., Chalmers D.C., Foster A.C., Vale W.W., Ling N. & Desouza E.B.: Displacement of insulin-like growth factors from their binding proteins as a potential treatment for stroke, *Proc. Nat. Acad. Sci. USA* **95** (1998) 1894–1898

Lozano G. & Montes de Oca Luna R.: MDM2 Function, *Biochim. Biophys. Acta* **1377** (1998) M55–M59

Martin J. L.; Baxter R.C.: IGF binding proteins as modulators of IGF actions, in: *The IGF system. Molecular biology, physiology, and clinical applications* (Eds. Rosenfeld R.G. & Roberts C.T.), Humana Press: Totowa, 1999 S. 227–255

McAlister M.S., Mott H.R., van der Merwe P.A., Campbell I.D., Davis S.J. & Driscoll P.C.: NMR analysis of interacting soluble forms of the cell-cell recognition molecules CD2 and CD48, *Biochemistry* **35** (1996) 5982–5991

Mendonça R., Palomares L. & Ramirez O.: An insight into insect cell metabolism through selective nutrient manipulation, *J. Biotechnol.* **72** (1999) 61–75

Midgley C.A. & Lane D.P.: p53 protein stability in tumor cells is not determined by mutation but is dependent on MDM2 binding, *Oncogene* **15** (1997) 1179–1189.

Momand J., Zambetti G.P., Olson D.C., George D. & Levine A.J.: The mdm-2 Oncogene Product Forms a Complex with the p53 Protein and Inhibits p53-mediated Transactivation, *Cell* **69** (1992) 1237–1245

Momand J., Jung D., Wilczynski S. & Niland J.: The MDM2 Gene Amplification Database, *Nucleic Acids Res.* **26** (1998) 3453–3459

Montelione G., Zheng D., Huang Y., Gunsalus K. & Szyperski T.: Protein NMR spectroscopy in structural genomics, *Nat. Struc. Biol.* **7** (2000) 982–985

Mori S., Abeygunawardana C., Johnson M. & van Zijl P.: Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation, *J Magn Reson B.* **108** (1995) 94–98

Mühlhahn P., Zweckstetter M., Georgescu J., Ciosto C., Renner C., Lanzenhöfer M., Lang K., Ambrosius D., Baier M., Kurth R. & Holak T.: Structure of interleukin 16 resembles a PDZ domain with an occluded peptide binding site, *Nat. Struc. Biol.* **5** (1998) 682–686

Neuhaus D. & Williamson M.: *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, Wiley-VCH, New York, 2. Aufl., (2000)

Noegel A., Rivero F., Albrecht R., Janssen K., Köhler J., Parent C. & Schleicher M.: Assessing the role of the ASP56/CAP homologue of Dictyostelium discoideum and the requirements for subcellular localization, *J.Cell Sci.* **112** (1999) 3195–3203

Öhman L., Ljunggren J. & Häggström L.: Induction of a metabolic switch in insect cells by substrate-limited fed batch cultures, *Appl. microbiol. Biotechnol.* **43** (1995) 1006–1013

Öhman L., Alacorn M., Ljunggren J., Ramqvist A. & Häggström L.: Glutamine is not an essential amino acid for Sf-9 insect cells, *Biotechnol. Lett.* **18** (1996) 765–770

Oliner J.D., Kinzler K.W., Meltzer P.S., George D.L. & Vogelstein B.: Wild-type p53 activates transcription in vitro, *Nature* **358** (1992) 80–83

Oliner J.D., Pietenpol J.A., Thiagalingam S., Gyuris J., Kinzler K.W. & Vogelstein B.: Oncoprotein MDM2 conceals the activation domain of tumor suppressor p53, *Nature* **362** (1993) 857–860

Ormoe M., Cubitt A., Kallio K., Gross L., Tsien R. & J. R.: Crystal structure of the *Aequorea victoria* green fluorescent protein, *Science* **273** (1996) 1392–1395

Park E.J., Park H.R., Lee J.S. & Kim J.: Licochalcone A - an inducer of cell differentiation and cytotoxic agent from pogostemon cablin, *Planta Med.* **64** (1998) 464–466

Pellecchia M., Sem D. & K. W.: NMR in drug discovery, *Nat. Rev. Dr. Disc.* **11** (2002) 211–219

Pellechia M., Sebbel P., Hermanns U., Wüthrich K. & Glockshuber R.: Pilus chaperone FimC-adhesin FimH interactions mapped by TROSY-NMR, *Nat. Struct. Biol.* **6** (1999) 336–339

Prendergast F.: Biophysics of the green fluorescent protein, *Methods in Cell Biology* **58** (1999) 1–18

Prestegard J., Valafar H., Glushka J. & Tian F.: Nuclear Magnetic Resonance in the Era of Structural Genomics, *Biochemistry* **40** (2001) 8677–8685

Rarey M., Kramer B., Lengauer T. & Klebe G.: A fast flexible docking method using an incremental construction algorithm, *J. Mol. Biol.* **261** (1996a) 470–489

Rarey M., Wefing S. & Lengauer T.: Placement of medium-sized molecular fragments into active sites of proteins, *J. Comput. Aided Mol. Design* **10** (1996b) 41–54

Reid D. (Ed.): *Protein NMR Techniques*, Humana, (1997)

Renner C. & Holak T.: NMR N-15 relaxation of the insulin-like growth factor (IGF)-binding domain of IGF binding protein-5 (IGFBP-5) determined free in solution and in complex with IGF-II, *Eur. J. Biochem.* **268** (2001) 1058–1065

Sali A.: 100,000 protein structures for the biologist, *Nat. Struct. Biol.* **5** (1998) 1029–1032

Salzmann M., Pervushin K., Wider G., Senn H. & Wüthrich K.: TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins, *Proc. Natl. Acad. Sci. USA* **95** (1998) 13585–13590

Seifert M.: *NMR-spektroskopische Studien zur strukturellen Dynamik des Grün-fluoreszierendes Proteins und der cAMP-abhängigen Protein Kinase*, Doktorarbeit, Technische Universität München, Fakultät für Chemie, (2002)

Seifert M., Ksiazek D., Azim K., Smialowski P., Budisa N. & Holak T.: Slow Exchange in the Chromophore of a Green Fluorescent Protein Variant, *J. Am. Chem. Soc.* **124** (2002) 7932–7942

Senn H., Eugster A., Otting G., Suter F. & Wüthrich K.: ^{15}N-labeled P22 c2 repressor for nuclear magnetic resonance studies of protein-DNA interactions, *Eur. Biophys. J.* **14** (1987) 301–306

Shibata S.: Shibata Lab. Nat. Med. Mater., *Minophagen Pharm. Co., Tokyo, Japan, Stem Cells (Dayton)* **12** (1994) 44–52.

Shuker S., Hajduk P., Meadows R. & Fesik S.: Discovering high-affinity ligands for proteins - SAR by NMR, *Science* **274** (1996) 1531–1534

Spyracopoulos L. & Sykes B.: Thermodynamic insights into proteins from NMR spin relaxation studies, *Curr. Opin. Struct. Biol.* **11** (2001) 555–559

Stoll R., Renner C., Mühlhahn P., Hansen S., Schumacher R., Hesse F., Kaluza B., Engh R., Voelter W. & Holak T.: Sequence-specific ^{1}H, ^{15}N, and ^{13}Cassignment of the N-terminal domain of the human oncoprotein MDM2 that binds to p53., *J. Biomol. NMR* **17** (2000) 91–92

Striker G., Subramaniam V., Seidel C. & Volkmer A.J.: Photochromicity and fluorescence lifetimes of green fluorescent protein, *J. Phys. Chem. B* **103** (1999) 8612–8617

Tabor S.: Expression using the T7 RNA polymerase/promotor system, in: *Current Protocols in Molecular Biology* (Eds. Ausubel F., Brent R., Kingsten R., Moore D., Seidman J., Smith J. & Struhl K.), Greene Publishing and Wiley Interscience, NY, 1990 S. 16.21–16.2.11

Tsien R.: The green fluorescent protein, *Annu. Rev. Biochem.* **67** (1998) 509–544

Ward W.: *Bioluminescence and Chemiluminescence: Basic Chemistry and Analytical Applications*, Academic Press, New York, 1981 S. 235–242

Wasylyk C., Salvi R., Argentini M., Dureuil C., Delumeau. I., Abecassis J., Debusche L. & Wasylyk B.: p53 mediated death of cells overexpressing MDM2 by an inhibitor of MDM2 interaction with p53, *Oncogene* **18** (1999) 1921–1934

Wattenberg L.: Chalcones, myo-inositol and other novel inhibitors of pulmonary cancerogenesis, *J. Cell. Biochem.* **Suppl. 22** (1995) 162–168

Wetterau L.A., Moore M.G., Lee K.W. & Shim M. L.and Cohen P.: Novel aspects of the insulin-like growth factor binding proteins, *Mol. Genet. Metabol.* **68** (1999) 161–181

Wijesinha-Bettoni R., Dobson C. & Redfield C.: Comparison of the Structural and Dynamical Properties of Holo and Apo Bovine α-Lactalbumin by NMR Spectroscopy, *J. Mol. Biol.* **307** (2001) 885–898

Wishart D., Sykes B. & Richards F.: Relationship between nuclear magnetic shift and protein secondary structure, *J. Mol. Biol.* **222** (1991) 311–333

Wolk A.: Can measurements of IGF-1 and IGFBP-3 improve the sensitivity of prostate-cancer screening, *Lancet* **356** (2000) 1902–1903

Wüthrich K.: *NMR of proteins and nucleix acids*, John Wiley & Sons, New York, (1986)

Yang F., Moss L. & Phillips Jr. G.: Molecular structure of green fluorescent protein, *Nature Biotech.* **14** (1996) 1246–1251

Yee A. et al.: An NMR approach to structural proteomics, *Proc. Natl. Acad. Sci. USA* (2002) 1825–1830

Zeslawski W., Beisel H.G., Kamionka M., Kalus W., Engh R.A., Huber R., Lang K. & Holak T.A.: The interaction of insulin-like growth factor-I with the N-terminal domain of IGFBP-5, *EMBO J.* **20** (2001) 3638–3644

Zhang O. & Forman-Kay J.: Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer, *Biochemistry* **34** (1995) 6784–6794

Zhu Z., Bulgakov O., S.S. S. & Dalton J.: Recombinant Expression and Purification of Human Androgen Receptor in a Baculovirus System, *Biochem. Biopys. Res. Comm.* **284** (2001) 828–835

Zurdo J., Guijarro J., Jimenez J., Saibil H. & C.M. D.: Dependence on Solution Conditions of Aggregation and Amyloid Formation by an SH3 Domain, *J. Mol. Biol.* **311** (2001) 325–340

Acknowledgements

I would like to thank Prof. Dr. Kessler for being my Doktorvater and Prof. Dr. Huber for the opportunity to work in his department.
I am thankful to Tad Holak for sharing his wisdom with me.
Furthermore I would like to thank all those colleagues and friends who accompanied me through this work:
The old masters Roland Baumgartner, Peter Mühlhahn, Christian Renner, Raphael Stoll and Markus Zweckstetter for teaching me the power of NMR.
Michael Brüggert, Julia Georgescu, Lars Israel, Mariusz Kamionka, Dorota Książek and Chrystelle Mavoungou for the good collaboration in various projects.
Ruth Pfänder and Markus Seifert for our work at the magnet and struggle with computers.
Conny Ciosto, Agata and Paweł Śmiałowski, Igor Siwanowicz and all the others of the NMR-group for a good working atmosphere.
Lazarus for Tuesday afternoons.
Johannes Wagner for continuous support.
Thomas Schlütsmeier for poetry and more.
Last but not least I'm indebted to my parents who supported me all along.