

Technische Universität München
Fakultät für Organische Chemie und Biochemie

Max-Planck-Institut für Biochemie
Abteilung Strukturforschung
Biologische NMR-Arbeitsgruppe

STRUCTURAL AND FUNCTIONAL STUDIES
ON PHOTOACTIVE PROTEINS AND
PROTEINS INVOLVED IN CELL
DIFFERENTIATION

Pawel Smialowski

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. Dr. A. Bacher

Prüfer der Dissertation:

1. apl. Prof. Dr. Dr. h.c. R. Huber

2. Univ.-Prof. Dr. W. Hiller

Die Dissertation wurde am 10.02.2004 bei der Technischen Universität
München eingereicht und durch die Fakultät für Chemie am 17.03.2004
angenommen.

PUBLICATIONS

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

Markus H. J. Seifert, Dorota Ksiazek, M. Kamran Azim, **Pawel Smialowski**, Nediljko Budisa and Tad A. Holak
Slow Exchange in the Chromophore of a Green Fluorescent Protein Variant
J. Am. Chem. Soc. 2002, *124*, 7932-7942.

Markus H. J. Seifert, Julia Georgescu, Dorota Ksiazek, **Pawel Smialowski**, Till Rehm, Boris Steipe and Tad A. Holak
Backbone Dynamics of Green Fluorescent Protein and the effect of Histidine 148 Substitution
Biochemistry. 2003 Mar 11; *42*(9): 2500-12.

Pawel Smialowski, Mahavir Singh, Aleksandra Mikolajka, Narashimsha Nalabothula, Sudipta Majumdar, Tad A. Holak
The human HLH proteins MyoD and Id-2 do not interact directly with either pRb or CDK6.
FEBS Letters (submitted) 2004.

ABBREVIATIONS

{I}-S NOE, nuclear Overhauser effect on nucleus S by saturating nucleus I

1D, one dimensional

2D, two dimensional

3D, three dimensional

4FW, 4-fluoro-tryptophan

5FW, 5-fluoro-tryptophan

6FW, 6-fluoro-tryptophan

CBCA(CO)NH, $^{13}\text{C}\beta$ - $^{13}\text{C}\alpha$ -(^{13}C)- ^{15}N - ^1HN correlation

CSA, chemical shift anisotropy

DD, dipole-dipole

f, NOE, enhancement factor

FID, free induction decay

GFP, green fluorescent protein

H-D, exchange, hydrogen-deuterium exchange

hetNOE, heteronuclear Overhauser effect

HNCA, ^1HN - ^{15}N - $^{13}\text{C}\alpha$ correlation

HNCO, ^1HN - ^{15}N - $^{13}\text{C}'$ correlation

HSQC, heteronuclear single-quantum coherence

kDa, Kilodalton

n, number of β -sheets in a β -barrel

NMR, nuclear magnetic resonance

NOE, nuclear Overhauser effect

NOESY, nuclear Overhauser effect spectroscopy

ppm, parts per million

R_1 , longitudinal relaxation rate

R_2 , transversal relaxation rate

S, shear number in β -barrels

T_1 , longitudinal relaxation time

T_2 , transversal relaxation time

TROSY, transverse relaxation-optimized spectroscopy

$\Delta\sigma$, anisotropy of the chemical shift tensor

$\Delta\nu$ NMR, line width

η , cross-correlation rate of ^{15}N -labeled CSA and ^1H - ^{15}N dipolar relaxation

τ_c , overall rotational correlation time

τ_m , mixing time

ω , Larmor frequency

INDEX

1	INTRODUCTION	1
1.1	Cell differentiation and cell cycle	1
1.1.1	Cell cycle	1
1.1.2	Interplay between cell differentiation and cell cycle.....	4
1.1.3	Helix-loop-helix (HLH) protein family	4
1.1.4	Id – 2	8
1.1.5	MyoD.....	8
1.1.6	CDK4/6.....	9
1.1.6.1	CDK4.....	9
1.1.6.2	CDK6.....	10
1.1.7	pRb.....	10
1.1.8	HPV16 virus E7 protein.....	11
1.2	Photoactive proteins.....	11
1.2.1	Function of GFP	11
1.2.2	α – PEC and photodynamic light sensing in plants.....	13
1.3	NMR (Nuclear Magnetic Resonance).....	15
1.3.1	Protein NMR	15
1.3.1.1	Principles of nuclear magnetic resonance	15
1.3.1.2	Larmor precession.....	17
1.3.1.3	Making Larmor precession of a nuclear spin observable	17
1.3.1.4	Energy levels, populations & signal-to-noise	17
1.3.1.5	Nuclear magnetic resonance	19
1.3.1.6	Relaxation and NMR linewidths	20
1.3.1.7	Chemical Shifts	20
1.3.1.8	J-coupling	21
1.3.1.9	Nuclear Overhauser effect (NOE).....	22
1.3.1.10	Exchange	23
1.3.1.11	Larger molecules, relaxation and TROSY	25
1.3.1.12	The NMR spectrum of a folded protein	27

2	MATERIALS AND METHODS	29
2.1	Materials.....	29
2.1.1	Chemicals	29
2.1.2	Enzymes and buffers	29
2.1.3	Kits and reagents	29
2.1.4	Oligonucleotides.....	29
2.1.4.1	Primers for MyoD mutagenesis.....	29
2.1.4.2	Primers for MyoD cloning to different vectors.....	30
2.1.4.3	Primers for MyoD gen synthesis	30
2.1.4.4	Primers for Id – 2 cloning.....	31
2.1.4.5	Primers for Id – 2 mutagenesis	31
2.1.4.6	Primers for E7 gene synthesis.....	32
2.1.4.7	Primers for E7 mutagenesis	33
2.1.4.8	Primers for GFP mutagenesis	33
2.1.5	Plasmids and constructs	33
2.1.5.1	Plasmids supplied by the companies	33
2.1.5.2	Plasmids constructed during the work.....	34
2.1.6	Bacterial strains	37
2.1.7	Peptides.....	38
2.1.8	Buffers and media	39
2.1.8.1	LB medium:	39
2.1.8.2	Minimal medium (MM) for uniform enrichment with ¹⁵ N	39
2.1.8.3	IPTG stock solution:	40
2.1.8.4	Antibiotics	40
2.1.9	Antibodies and other proteins	41
2.1.9.1	Antibodies	41
2.1.9.2	Molecular weight marker for SDS-PAGE electrophoresis.....	41
2.1.10	Other chemicals.....	41
2.1.10.1	Protease inhibitors:.....	41
2.1.10.2	Isotopically enriched chemicals:	42
2.1.10.3	Miscellaneous :	42
2.2	Apparatus	42

2.2.1	ÄKTA explorer 10 purification system	42
2.2.1.1	Chromatography equipment, columns and media:	43
2.2.2	NMR spectrometers.....	43
2.2.3	Other apparatus.....	43
2.3	Consumables	44
2.4	Methods.....	45
2.4.1	DNA techniques	45
2.4.1.1	Basic cloning techniques	45
2.4.1.2	Isolation of the plasmids	45
2.4.1.3	Screening of positive colonies	45
2.4.1.4	Restriction assays	45
2.4.1.5	DNA sequencing	46
2.4.1.6	Cloning to LIC vectors	46
2.4.1.7	Gen synthesis	48
2.4.1.8	Site directed mutagenesis.....	48
2.4.1.9	In vitro protein synthesis.....	50
2.4.1.10	Escherichia Coli transformation	51
2.4.2	Bacterial culturing	53
2.4.2.1	Bacterial culture in LB medium:	53
2.4.2.2	Bacterial culture in MM:.....	53
2.4.2.3	Medium for selectively enrichment amino acids	54
2.4.3	Protein production techniques.....	55
2.4.3.1	Protein expression in Escherichia Coli strains	55
2.4.3.2	Expression in Spodoptera frugiperda strain SF9.....	56
2.4.4	Protein purification techniques	56
2.4.4.1	Cell differentiation and cell cycle proteins.....	56
2.4.4.2	GFPuv and mutant proteins	57
2.4.5	Handling and storing of the proteins.....	58
2.4.6	Analytical methods.....	58
2.4.6.1	Protein detection	58
2.4.6.2	Pull down assays	61
2.4.6.3	Analytical gel filtration	61

2.4.6.4	Mass spectroscopy	61
2.4.6.5	Functional assays of cell cycle and differentiation proteins	61
2.4.6.6	Optical Spectroscopy	62
2.4.6.7	NMR	62
2.4.6.8	NMR diffusion measurements	63
2.4.6.9	NMR relaxation measurements	64
2.4.6.10	Other methods	65
2.4.7	Bioinformatics und calculations	65
3	RESULTS AND DISCUSSION.....	66
3.1	Cell differentiation and cell cycle proteins	66
3.1.1	The HLH domain of MyoD or Id-2 does not interact with the pocket domain of pRb	66
3.1.1.1	Results	66
3.1.1.2	Discussion	72
3.1.2	There is no interaction between human or chicken MyoD and CDK6	73
3.1.2.1	Results	73
3.1.2.2	Discussion	75
3.2	Photodynamic proteins	76
3.2.1	GFPuv	76
3.2.1.1	Results	76
3.2.1.2	Discussion	90
3.2.2	¹⁹ F fluoro tryptophane GFP.....	96
3.2.2.1	Results	96
3.2.2.2	Discussion	102
3.2.3	α -PEC (α - subunit of phycoerythrocyanin)	107
3.2.3.1	Results	107
3.2.3.2	Discussion	114
4	SUMMARY	115
5	ZUSAMMENFASSUNG	117
6	REFERENCES.....	119
7	SUPPLEMENTARY MATERIALS.....	134

7.1	CDK4/6 structurally guided alignment	134
7.2	Bioinformatics und calculations	136

1 INTRODUCTION

1.1 Cell differentiation and cell cycle

1.1.1 *Cell cycle*

The cell cycling process is carefully regulated and responds to the specific needs of a certain tissue or cell type. Normally, in adult tissue, there is a delicate balance between cell death (programmed cell death or apoptosis) and proliferation (cell division) producing a steady state. Disruption of this equilibrium by loss of cell cycle control may eventually lead to tumor development (Sherr 1996).

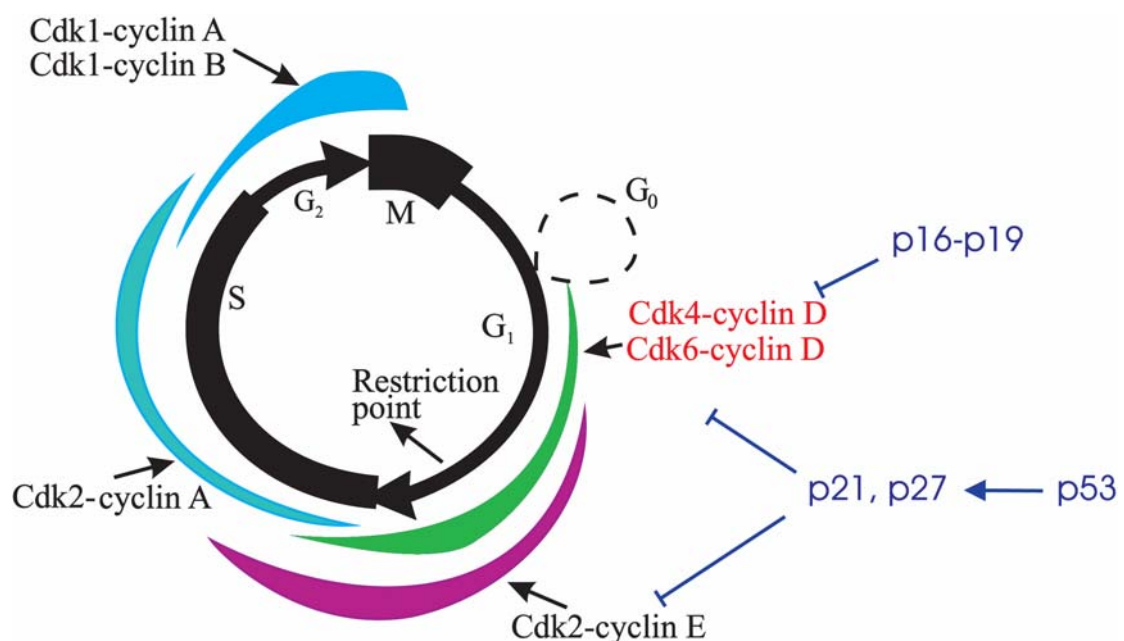


Figure 1.1. During cell cycle precisely regulated expression of CDK's (cyclin dependent kinases), cyclins, and CDK's inhibitors take place. The lines thickness corresponds to the level of activity of given kinases.

The highly organized and regulated cell cycle process is responsible for duplication of the cell (Figure 1.1 and Figure 1.2). Tight regulation and timing ensure that DNA is replicated once during the S phase (without errors), and that identical chromosomes are equally delivered to daughter cells during the M phase (Heichman et al. 1994; Sherr 1996).

The cell cycle is, therefore, an alteration of two main processes: A) the "doubling" process (S = synthesis phase) where DNA is synthesized, and B) the "halving" process (M = mitosis phase) where the cell and its contents are divided equally into two daughter cells. The periods between these processes are called gap periods (G phase). Taken together, the cell cycle consists of the different phases listed in table 1.1 (Sandal et al. 2002).

Phase	Description
G ₁	Growth and preparation of the chromosomes for replication
S	Synthesis of DNA (centrioles/DNA replication)
G ₂	Preparation for mitosis
M	Mitosis
G ₀ *	Temporary or permanent state of cell cycle exit. Postmitotic/terminally differentiated

*Exit from the cell cycle at G₁, not occurring in every cell cycle.

Table 1.1. Cell cycle phases.

The switch, or transition, between phases is a hallmark of the cell cycle, with an extremely accurate timing and order of molecular events. However, if something goes wrong, the cell has several systems for interrupting the cell cycle. These are the quality control points of the cell cycle and are often referred to as checkpoints (Elledge 1996). At checkpoints, there are important mechanisms sensing damaged DNA before the cell enters the S phase (G₁ checkpoint) or the M phase (G₂ checkpoint) (Figure 1.1). One major molecular hallmark of checkpoint control is where transitions turn off the previous state and promote the future state of the cell cycle (irreversible progression). Loss of checkpoint control results in genomic instability, accumulation of DNA damage, uncontrolled cell proliferation, and, eventually, tumorigenesis. Indeed, this has been implicated in the progression of many human cancers (Elledge 1996; Sherr 2000).

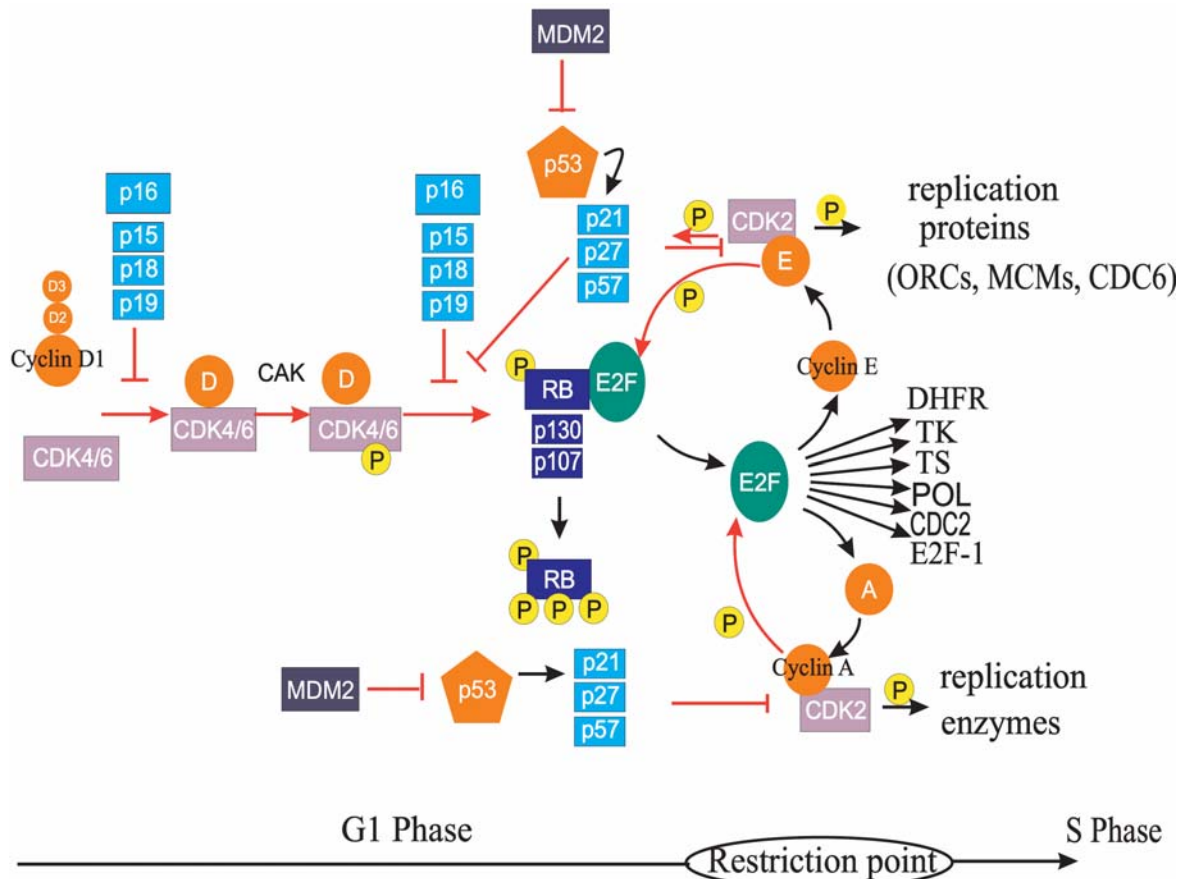


Figure 1.2. Network of protein – protein interactions on the onset of G1 phase of cell cycle serve as a precise switch. If cell is stimulated by myogenic signals from environment it allows passing through G1 restriction point. It accumulates outside signals and cause expression of proteins indispensable for DNA synthesis through S phase. Function and mode of action of selected proteins involved in the cell cycle regulation are described in text.

The G₁-S transition is a highly regulated and important transition in the cell cycle. At this stage, the cell cycle passes a point between the G₁ and S phase (restriction point) with an irreversible commitment to a new cycle. Stimulation of cells with mitogens results in cell cycle progression. G₁ progression requires sustained cyclin D expression, persisting as long as there are ongoing mitogenic signals. Mitogenic transcription activation of D-type cyclins has been shown to be induced, for instance, by c-Myc, AP-1, and NF- κ B transcription factors (Ekholm et al. 2000). However, when mitogens are removed, the level of cyclin D rapidly decreases and the cell is arrested in G₁. A similar block in cyclin D-kinase activity and subsequent arrest in G₁ phase is achieved by the INK4a family of CKIs (Figure 1.2) (Sherr 1996). Four different members of this family (p16INK4a, p15INK4a, p18INK4a, and

p19INK4a) are known to bind and inhibit CDK4 and CDK6, without affecting other CDKs (Sherr 1996). The INK4a-bound CDKs are not able to complex with cyclin, and stay intact as an INK-CDK heterodimer. In contrast, the Cip/Kip family of CKIs (p27Cip/Kip, p21Cip/Kip, and p57Cip/Kip), which do not inhibit the G1 phase, play a positive role by stabilizing the CDK/cyclin complex (Cavenee et al. 1995; Sherr et al. 1999). Cyclins are unstable proteins and their levels vary throughout the cell cycle. This is due to degradation by the ubiquitin/proteasome pathway when they are not required. CDK activity is inhibited by phosphorylation on specific tyrosine residues, and phosphatase treatment leads to a hyperactive kinase (Heichman et al. 1994). Three different mammalian phosphatases are known, Cdc25 A, B, and C. The regulation of Cdc25A is critical for G₁ response to DNA damage. The main role of the CDK4/6 / cyclin D complex in the early progression of G1 is to phosphorylate pRb and thereby promote cell cycle progression. The pRb phosphorylation releases E2F and allows the expression of regulators required for DNA synthesis and S phase progression. E2F triggers expression of proteins like dihydrofolate reductase, thymidine kinase, different DNA polymerases and the late-G₁ cyclin E (Renan 1993). Expression of cyclin E establishes a positive feedback loop of pRb phosphorylation, since cyclin E in complex with CDK2 will continue to phosphorylate pRb (Bishop 1996), contributing to an irreversible transition into the S phase and cell cycle progression, even in the absence of growth factors (Harbour et al. 2000). The biochemical events are primarily phosphorylation, dephosphorylation, and ubiquitination, with the overall mission to either prevent or induce a new cell cycle via the pRb pathway.

1.1.2 Interplay between cell differentiation and cell cycle

Cell differentiation is crucial for growth and function of any multicellular organism. In order to execute cell differentiation on the cellular level there has to be interplay between differentiation signals and the cell cycle. The proteins from the HLH family are crucial for coordination of the differentiation processes with the cell cycle. Large body of evidence indicates that protein – protein interactions are important in pathways leading to and from cell cycle.

1.1.3 Helix-loop-helix (HLH) protein family

There are two main categories of the basic helix-loop-helix (bHLH) proteins. The class A bHLH, also known as the E proteins, such as those encoded by differentially spliced

transcripts from the E2A (E12, E47 and E2-5/ITF1 proteins), E2-2/ITF2, and HEB/HTF genes, are ubiquitously expressed (Murre et al. 1989; Henthorn et al. 1990; Massari et al. 2000). The class B bHLH proteins, which include members such as MyoD, myogenin, NeuroD/BETA2, and TAL, show a tissue-restricted pattern of expression. Dimerization is essential for DNA binding and transcriptional activity *in vivo* (Lassar et al. 1991; Massari et al. 2000) and in general, class B bHLH proteins form heterodimers with the class A bHLH proteins, although the latter can also operate as homodimers (Shen et al. 1995). The basic region of each protein is required for binding to DNA, commonly to a region that includes a specific sequence motif known as the E-box (CANNTG) (Ephrussi et al. 1985; Kiledjian et al. 1988; Lassar et al. 1989) or the related N-box (CACNAG) (Klamt et al. 1989; Tietze et al. 1992). Basic informations about biological functions of members of both class-A and class-B are listed in the table 1.2.

Id is another category of mammalian HLH protein, which was first identified by Benezera et al. in 1990. So far four members of the Id family, Id-1 to Id-4, have been identified (Benezera et al. 1990; Christy et al. 1991; Pagliuca et al. 1995). Id proteins bind to both class-A and class-B bHLH proteins and inhibit their ability to bind DNA as homodimeric or heterodimeric complexes. Because of this activity, these proteins were named as Id (inhibitor of DNA binding).

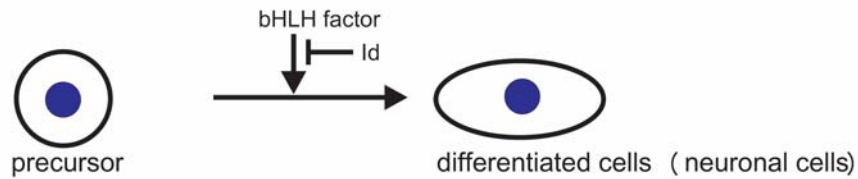
During development, all members of the Id family are expressed in many organs and tissues, showing overlapping but distinct expression patterns in developmental stage-dependent and region-specific manners (Riechmann et al. 1995; Jen et al. 1996; Jen et al. 1997). As demonstrated by *in situ* hybridization studies, a general feature of Id gene expression during development is that the distributions of Id-1, Id-2 and Id-3 are overlapping, while Id-4 shows a unique pattern during embryogenesis (Riechmann et al. 1995; Jen et al. 1996; Jen et al. 1997). It is clear that inhibitory and stimulatory factors are important in achieving a correct body plan. Id proteins are typically not expressed in most mature tissues (Biggs et al. 1992; Ellmeier et al. 1992). Figure 1.3 sum up functions of Id proteins.

	Protein	Source	Biological function
Class A	E12, E47, E2-5	Mammalian	Generated by alternative splicing of E2A gene. Known collectively as E-proteins. They form homodimers and heterodimers with class B bHLH proteins. Required for lymphocytes B differentiation
	Daughterless	Drosophila	Functional homologue of mammalian proteins.
Class B	Myogenic		
	MyoD, Myf-5, MRF4, Myogenin	Mammalian	Initiates the myogenic programme in many cell types. Mice lacking MyoD and Myf-5 are devoid of skeletal myoblasts and muscle. MRF4 and myogenin are expressed during muscle cell terminal differentiation to regulate myotube formation and fusion.
	Neurogenic		
	Neurogenin, NeuroD1, NeuroD2, NeuroD3	Mammalian	Involved in terminal neuronal differentiation. Ectopic expression activates neuron-specific genes and initiates the neural differentiation programme in non-neural cells.
	NSCL-1, NSCL-2	Mammalian	Expression pattern restricted to the nervous system.
	Mash-1, Mash-2	Mammalian	Neural and trophoblast development
	Hes-1,2,3,5	Mammalian	Hes-1, 2 widely expressed in embryonic and adult tissues. They repress promoters containing both E- and N-box motifs. Hes-3 expression is restricted to cerebellar Purkinje cells and Hes-5 to developing nervous system.
	Haemopoietic		
	SCL, Lyl-1	Mammalian	Required for haemopoiesis. Represses E-box-dependent transcription by E47.
	Cardiogenic		
eHand, dHand	Mammalian	Regulate morphogenetic events of asymmetric heart development. eHand expression is regulated by the homeodomain factor Nkx2-5	

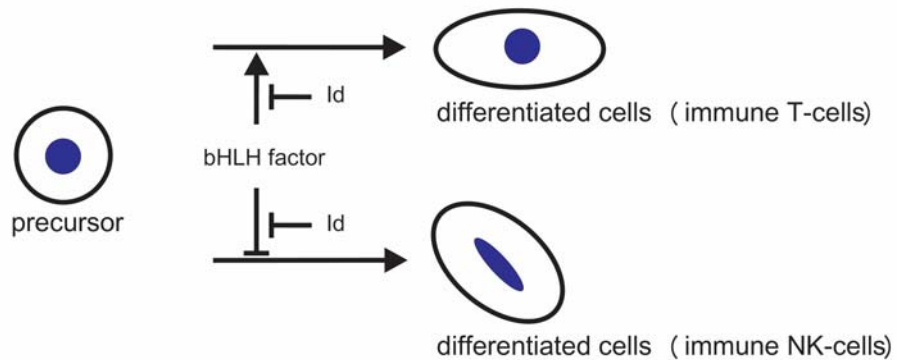
Table 1.2. Name, functions and natural occurrence of some of bHLH proteins.

A. Differentiation control by Id

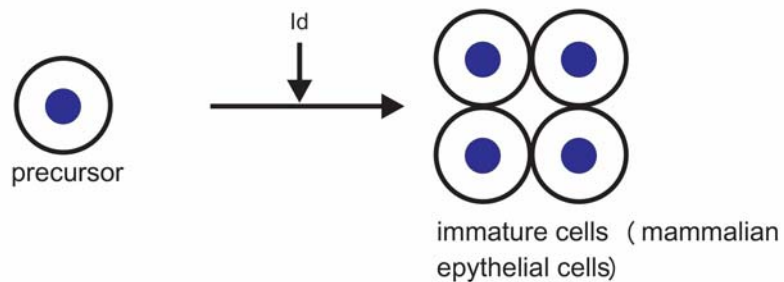
1) One-way type (stops differentiation)



2) Bifurcation type (change cell fate by changing maturation pathway)



B. Proliferation control by Id



C. Survival promotion by Id

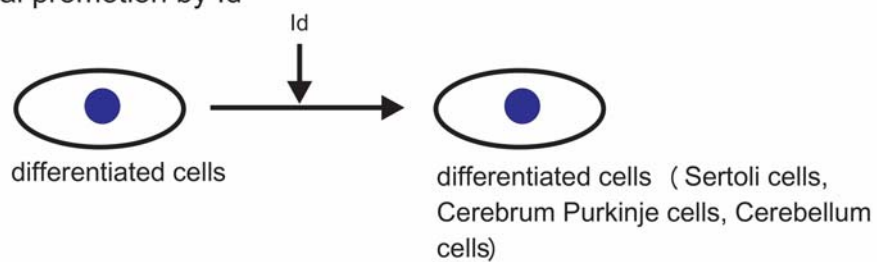


Figure 1.3. The different function of Id proteins in differentiation, proliferation and survival processes. (A) Differentiation controlled by bHLH factors could be influenced by Id either by holding cells in undifferentiated state or by switching to distinct pathway. A bHLH factor promotes cell differentiation and Id inhibits it by suppressing the function of the bHLH factor. The bifurcation type is seen in the case of NK cell development. Id is an inhibitor for T - cell but a stimulator for NK - cell. (B) Id maintains the immaturity of cells and stimulates their proliferation. (C) Id functions as a survival factor for postmitotic differentiated cells.

1.1.4 Id-2

Only Id-2, and not the other members of the family, unlike Id-1 and Id-3, it is able to disrupt the antiproliferative effects of tumor suppressor proteins of the pRb family (the 'pocket' proteins: pRb, p107 and p130), thus allowing cell-cycle progression (Sandal 2002). Although the HLH region is well conserved among Id family members, only the HLH domain of Id-2 was shown to interact with pRb, p107, p130 (Iavarone et al. 1994; Lasorella et al. 1996). It is claimed that GST-Id-2 and GST-MyoD binds with comparable efficiency to wild type pRb (Lasorella et al. 1996).

Correct expression of Id-2 is essential to regulate proliferation and differentiation of the neural crest, thus neural crest precursor cells might be sensitive to inappropriate expression of Id-2 (Jogi et al. 2002). Id-2 construct lacking HLH domain did not modify the block of progression through S phase caused by pRb (Iavarone et al. 1994). Binding of both E1A and Id-2 involves similar region in pRb, which must retain an intact pocket domain (Iavarone et al. 1994; Lasorella et al. 2000).

1.1.5 MyoD

The MyoD is a muscle determination factor (MDF) belonging to a group of muscle-specific basic helix-loop-helix (bHLH) transcription factors. In a complex regulatory network with myf-5, myogenin, and MRF4, MyoD plays a key role in the determination and differentiation of all skeletal muscle lineages (Weintraub et al. 1991; Arnold et al. 1996; Arnold et al. 1998). Following isolation of the MyoD gene by Davis et al. (1987), it was shown that the transfection and subsequent forced expression of the MyoD gene in a variety of differentiated non-muscle cell types (fibroblast, fat, melanoma, neuroblastoma, chondroblast, liver, and retinal pigmented epithelial cell lines) can initiate the process of myogenesis (Davis 1987; Weintraub 1989; Choi 1990). Forced expression of MyoD in a variety of different cell backgrounds, including normal, transformed and tumor cell lines, was also shown to induce growth arrest. This feature was correlated with the presence of the MyoD bHLH domain (Crescenzi et al. 1990).

Growth arrest required functional pRb (Schneider et al. 1994; Novitch et al. 1996). *In vitro* protein binding and immunoprecipitation studies indicated that both MyoD and myogenin bind to hypophosphorylated pRb directly through the bHLH domain (Gu et al. 1993). There has been however discrepancy among experimental reports about the MyoD

and the pRb interaction. There are some publications reporting both *in vivo* and *in vitro* interactions (Gu et al. 1993), other showing that there is no interaction whatsoever (Zhang et al. 1999; Zhang et al. 1999; Puri et al. 2001).

In addition to the interaction of the bHLH domain of MyoD was shown recently (Zhang et al. 1999; Zhang et al. 1999) to interact with CDK4 through a conserved 15 amino acid region situated at the C-terminus of MyoD. The C-terminus of MyoD or 15 amino acid peptide was shown to be sufficient to inhibit CDK4-dependent phosphorylation of pRb. It was further claimed that CDK4 disrupts the DNA-binding and transactivation functions of MyoD through the formation of the MyoD-CDK4 complex that does not required kinase activity (Zhang et al. 1999; Zhang et al. 1999) A general model emerged suggesting that pRb mediates terminal cell cycle withdrawal during muscle cell differentiation by interacting directly with the myogenic factors to block pRb phosphorylation and cell growth.

It was also shown (Iavarone et al. 1994) that GST-Id-2 and GST-MyoD bind with comparable efficiency to the wild type pRb and that there is no binding found for the pRb 706 (C->F) mutant. Point mutation in pRb 706 (C->F) abolish also binding of viral proteins SV40 large T antigen and E1A to pRb.

1.1.6 CDK4/6

1.1.6.1 CDK4

Harbour et al. (1999) presented evidence that phosphorylation of the C-terminal region of pRb by CDK4/CDK6 initiates successive intramolecular interactions between the pRb C-terminal region and its central pocket. This initial interaction displaces histone deacetylase from the pocket, blocking active transcriptional repression by pRb. This facilitates a second interaction that leads to phosphorylation of the pocket by CDK2 (Sandal 2002) and disruption of the pocket structure. These intramolecular interactions provide a molecular basis for sequential phosphorylation of pRb by CDK4/CDK6 and CDK2. CDK4/CDK6 is activated early in G1, blocking active repression by pRb. However, it is not until near the end of G1, when cyclin E (Harbour et al. 1999) is expressed and CDK2 is activated, that pRb is prevented from binding and inactivating E2F (Knudsen et al. 1997).

1.1.6.2 *CDK6*

CDK6 shows a considerable amino acid homology with CDK4, and the 2 proteins possess very similar 3D structures (Supplementary materials) (Guex et al. 1997; Schwede et al. 2003). Functionally, both proteins bind cyclin D1 and are specifically inhibited by CDKN2A (Serrano et al. 1993).

1.1.7 *pRb*

pRb inhibits cell cycle progression by interacting with transcription factors such as E2F; when pRb becomes phosphorylated, E2F is released and stimulates proliferation (Lee et al. 1987; Lee et al. 1987; Knudsen et al. 1997; Harbour et al. 1999). Besides cell cycle inhibition through E2F suppression, pRb has also been shown to suppress apoptosis. For instance, pRb-deficient embryos show defects in fetal liver haematopoiesis, neurogenesis and lens development, and extensive apoptosis was observed in these tissues (Morgenbesser et al. 1994; Macleod et al. 1996). The mechanisms by which pRb/E2F influences apoptosis remain unknown. E2F has been shown to induce the expression of the pro-apoptotic factor Apaf-1 and evidence suggests a role for E2F in apoptosis following DNA damage (Blattner et al. 1999; Moroni et al. 2001). E2F cannot induce apoptosis when pRb is co-expressed and pRb possibly has an anti-apoptotic effect through the inhibition of E2F (Fan et al. 1996; Pucci et al. 2000).

p53 and pRb/E2F may be directly linked in cell proliferation and apoptosis (Fan et al. 1996; Pucci et al. 2000). Activated p53 causes a G1 arrest by inducing p21, followed by an inhibition of cyclin/CDK (Sandal 2002). In these conditions, pRb is not phosphorylated and cells do not progress through the cell cycle (Harbour et al. 1999). In contrast, free E2F directly induces p53 transcription, thus connecting the pRb/E2F pathway to p53-dependent apoptosis (Hiebert et al. 1995). Each of both tumor suppressors (p53 and pRb) may thus be able to compensate for the loss of the other (King et al. 1998).

Three DNA tumor virus-transforming proteins, including SV40 T antigen, adenovirus E1A, and human papilloma virus E7 bind to pRb and reverse its growth suppressive function (DeCaprio et al. 1988; Whyte et al. 1988). Only the unphosphorylated form of pRb binds to these viral oncoproteins (Weinberg 1991).

1.1.8 HPV16 virus E7 protein

Human Papilloma Virus (HPV) is a sexually transmitted disease affecting primarily the cervix, but also the male genitalia. It is more common in anemic younger women (>30 years old), contraceptive hormone users, smokers, the poor, during late pregnancy, and in periods of low immunity (Santos et al. 2001; Moreno et al. 2002).

Human papilloma viruses (HPV) are a family of DNA viruses that can infect, prolong the life, and transform epithelial cells, particularly in the skin, head, neck and, cervix (Morozov et al. 1997; Enzenauer et al. 1998). The expression of multiple viral proteins is mandatory for rapid and complete transformation. The E7 protein is the major oncogenic protein produced by cervical cancer-associated HPV16 (Enzenauer et al. 1998). The majority of the transforming potential of E7 protein has been ascribed to its ability to bind to, and down-regulate, the retinoblastoma (RB) gene product (Alesse et al. 1998; Lochmuller et al. 1999). Expression of E7 proteins can also increase protein levels of p21 (Morozov et al. 1997; Alesse et al. 1998). Depending upon the degree of E7 expression and the differentiation state of the cell type, increased p21 levels can either enhance DNA synthesis and apoptosis or cause growth arrest (Morozov et al. 1997).

1.2 Photoactive proteins

1.2.1 Function of GFP

Green fluorescent proteins (GFP's) have typical β -can fold consisting of 11 β -sheets closed around central α – helix (Figure 1.4). They provide powerful tools used in molecular biology for monitoring gene expression, protein movement and protein interactions (Philips 1997; Tsien 1998; Prendergast 1999) and also exhibit complex photophysical behavior as observed by ensemble and single molecule measurements (Tsien 1998; Prendergast 1999). The optical absorption, fluorescence and ultra-fast dynamic fluorescence properties of GFP are commonly explained by a three-state model (Philips 1997; Tsien 1998; Prendergast 1999) which assumes that the chromophore exists either in a neutral form (A) (395 nm) or an anionic (475 nm) state; the latter state exists in a thermodynamically unstable intermediate form (I) (493 nm) and a stable low energy form (B). Fluorescence dynamics are rationalized by the model of an intramolecular Förster cycle (Föster 1950) between the states, i.e., $A \rightarrow A^* \rightarrow I^* \rightarrow I \rightarrow A$, where the asterisks indicate excited states; A – ground

state; I – intermediate state (Föster 1950; Tsien 1998). Upon illumination of GFP with UV light, an $A \rightarrow B$ photoconversion takes place that reduces the absorbance of state A and increases the absorbance of state B (van Thor et al. 1998; van Thor et al. 2002).

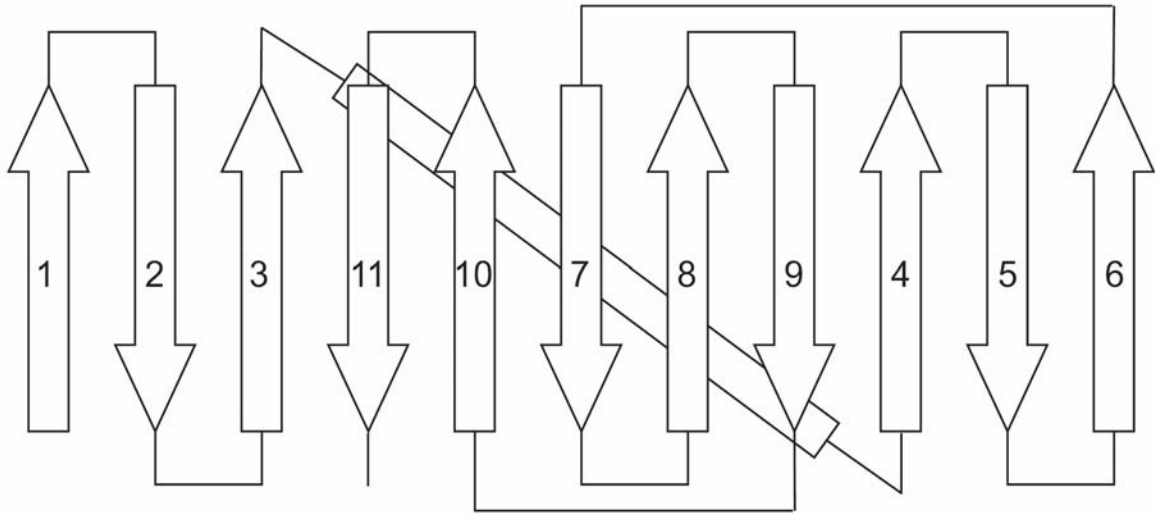


Figure 1.4. Secondary structure elements of GFP. Numbered arrows correspond to β – sheets. Long box at the background correspond to internal α – helix which contain chromophore.

Other models put forward assumption that cationic forms, in which the nitrogen of the imidazole ring of the chromophore is protonated, are responsible for the absorbance spectra (Voityuk et al. 1997; Voityuk et al. 1998). The model of an intramolecular Förster-cycle (Voityuk et al. 1997) rationalizes the fluorescence dynamics. GFPs exhibit a variety of photochemical processes including photoconversion (van Thor et al. 2002) and complex on-off blinking and switching behavior in single-molecule measurements, which cannot be easily explained on the basis of the canonical three-state model (Weber et al. 1999).

Up to now the family of GFP-like proteins comprise 27 cloned and spectroscopically characterized proteins of which 20 high-resolution crystal structures are available. The overall structure of GFP consists of an 11-stranded β -barrel with a central helix that carries the chromophore (Philips 1997). 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.

Competition of radiationless photoisomerization and fluorescence emission has been suggested to rationalize the differences in quantum yield of fluorescence in various GFP

mutants. This is underlined by the fact that at room temperature denatured GFP or model chromophores in solution show no significant fluorescence. However, when frozen as ethanol glass at 77°K, these compounds become highly fluorescent. Inside the protein environment the chromophore is very sensitive to mutations of surrounding residues that alter the permitted degrees of freedom of the chromophore (Kummer et al. 2000). Structural rigidity in combination with the high melting temperature of *Aequorea* GFP [78°C (Ward 1981)] and the stability of the GFP fold against denaturants and proteases lead to the view of GFP being a very stable and inflexible protein (Ward 1981; Tsien 1998). It has been shown that the fluorescence properties of GFP, as well as red fluorescent variant dsRed, and of the chromoprotein asCP are very sensitive to mutations of residues 148, 94 and 203 which are supposed to play a major role for proton transfer during the intramolecular Förster cycle. The NMR studies on GFPs turned out to be a challenge due to the strong tendency of GFPs to aggregate (Seifert et al. 2002).

1.2.2 α – PEC and photodynamic light sensing in plants

Phycobilisomes, the light-harvesting systems of cyanobacteria, red algae, and cryptomonad algae, are composed of phycobiliproteins and linker proteins (Glazer 1985). Phycobiliproteins are classified into four main groups: phycoerythrin (PE), phycocyanin (PC), phycoerythrocyanin (PEC), and allophycocyanin (APC), depending on their absorption maxima. They are all similar in that their monomeric units consist of two subunits, called a and b, which are highly homologous to each other, and are also highly homologous between different types of proteins (Glazer 1985; Beatz 1997).

The phycoerythrocyanin (PEC) complexes are located at the periphery of phycobilisomes (Bryant 1982). PEC of the thermophilic cyanobacterium *Mastigocladus laminosus* is the best characterized complex of this biliprotein class (Reuter et al. 1993). The α -subunit has a single phycoviolobilin (PVB) chromophore covalently linked via a thioether bond to Cys-84, whereas the β -subunit contains two phycocyanobilin (PCB) chromophores (Duerring et al. 1990). Unlike other phycobiliproteins, the phycoviolobilin chromophores of the PEC-“trimers” present a remarkable reversible photochemistry (Björn 1979). This finding initiated intense investigations on this unusual spectroscopic behavior especially regarding the possible function as a sensor pigment (Hong et al. 1993).

It should be remembered that chromoproteins (like for example, α -PEC and GFP) reduce the conformational freedom of the chromophores. The chromophores are non-fluorescent and photochemically inactive in the free state or in the denatured protein, but due to the conformational restrictions, become fluorescent and photochemically active in the native protein.

Biochemical and spectral data assign the photochemistry exclusively to the α -PEC subunit (Bishop et al. 1987). Similar to phytochrome and phytochrome-like photoreceptors of higher plants and cyanobacteria, the PVB-chromophore undergoes spectral and molecular changes that are dependent on the light-induced state (Neff et al. 2000). The phototransformation of the α -PEC chromophore is reflected by a reversible shift in the visible absorption maximum from 505 to 570 nm and the two states were termed E and Z, respectively. Isomerization can be performed by irradiation with complementary chromatic light and the two states are stable in the dark. The molecular mechanism of the photoreaction in α -PEC is expected to be similar to that of the phytochromes and would involve isomerization of the chromophore around the D15, 16 double bond between the pyrrole rings C and D (Zhao et al. 1995; Neff et al. 2000). In addition, however, at least two types of photochemistry, type I and type II Z/E-isomerizations, have been spectroscopically observed in α -PEC (Figure 1.5).

The ratio of the two types is believed to be controlled by sulfhydryls groups of Cys-98 and Cys-99 in the protein (Hong et al. 1993). Whereas the structures of the PVB isomers are well defined by spectroscopic characterizations, the participation of the protein in the phototransformation process of the α -subunit is nearly unknown.

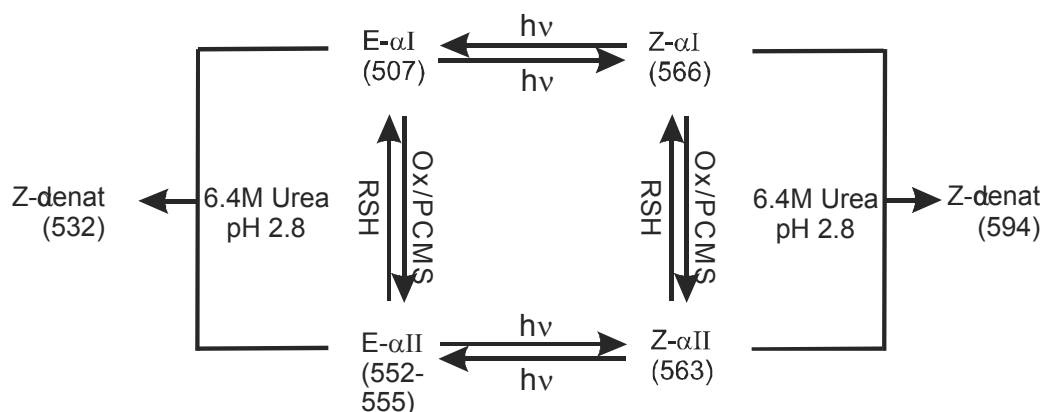


Figure 1.5. Type I and type II photochemistry of a-PEC, see reference (Zhao, K.-H., et al. 1995) for details.

The aim in this part of my research was to study the nature of the chromophore – protein interaction in the photoactive proteins. These types of interactions contribute to both energy transfers like for example, in the photosystem I and II in plants or for light sensing in the phytochrome.

1.3 NMR (Nuclear Magnetic Resonance)

1.3.1 *Protein NMR*

Nuclear magnetic resonance spectroscopy plays an important role in structural investigations complementing X-ray crystallography for small and medium size proteins (below 30 kDa) (Montelione et al. 2000; Prestegard et al. 2001). NMR can deliver strong results in several areas of structural biochemistry. It is the basis for a wide range of experiments and can be used to determine structure-function relationships (Shuker et al. 1996), to find binding partners with their specific binding sites (Stoll et al. 2001), to investigate dynamics of proteins, to distinguish multiple conformations (Muhlhahn 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 2003). A series of spectra taken under different conditions may be used to monitor aggregation and 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 weak ligands binding to target molecules has made NMR also increasingly important in drug discovery (Diercks et al. 2001; Pellecchia et al. 2002).

1.3.1.1 *Principles of nuclear magnetic resonance*

The nuclear spin that gives rise to a magnetic dipole moment $\mu = \gamma I$, depends on the atom type:

Nucleus	I	γ (T*s) ⁻¹	γ (relative)	Natural abundance (%)
1H	1/2	2.675*108	1	99.98
2H	1	4.107*107	0.15	0.02
13C	1/2	6.728*107	0.25	1.11
15N	1/2	-2.712*107	0.1	0.36
31P	1/2	1.084*108	0.41	100

Table 1.3. Basic magnetic properties of biochemically relevant nuclei.

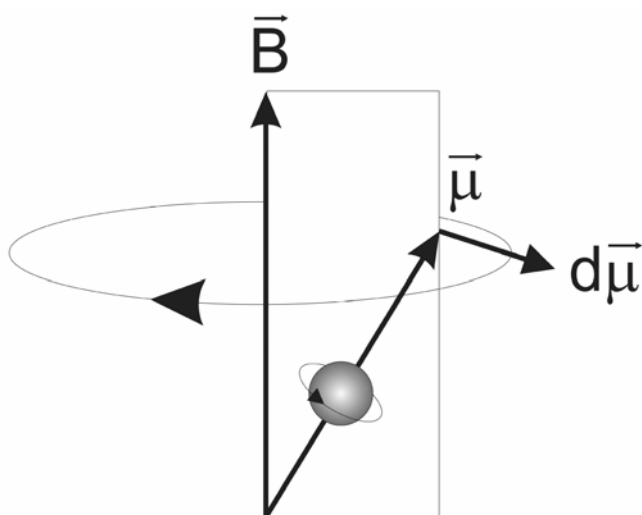


Figure 1.6. The spinning nucleus with a charge precessing in a magnetic field.

The energy associated with a magnetic dipole moment in a magnetic field of strength B_0 is given by $E = -\mu B_0 = \gamma I B_0$. The nuclear spin moments are quantified, thus, only one of the three Cartesian coordinates, I_z can be specified simultaneously with the magnitude of the spin, I^2 . For a spin $\frac{1}{2}$ nucleus, two $(2I + 1)$ orientations are possible ($I_z = \pm \frac{1}{2}$), and the corresponding energy levels are populated according to the Boltzmann distribution.

1.3.1.2 *Larmor precession*

When a nucleus with a nuclear magnetic moment is placed in an external magnetic field B , (Figure 1.6), the magnetic field of the nucleus will not simply be oriented in the direction of the magnetic field, but because the nucleus is rotating, the nuclear magnetic field will instead precess around the axis of the external field vector. This is called *Larmor precession*. The frequency of this precession is a physical property of the nucleus and it is proportional to the strength of the external magnetic field, the higher the external magnetic field the higher the frequency (Figure 1.8).

1.3.1.3 *Making Larmor precession of a nuclear spin observable*

The magnetic moment of a single nucleus cannot be observed. In a sample with many nuclei the magnetic moment of the individual nuclei will add up to one component. In a sample of nuclei in a magnetic field the component of the magnetic moment of the precessing nuclei around the external field axis will be a vector in the direction of the field axis. This nuclear magnetization is impossible to observe directly. In order to observe the nuclear magnetization we want to bring the nuclear magnetization perpendicular to the applied field. Applying a radio frequency pulse, which is perpendicular to the external magnetic field, can do this. If this pulse has the same frequency as the Larmor frequency of the nuclei to be observed and of a well defined length, then the component of the nuclear magnetization can be directed away from the direction of the magnetic field to a direction perpendicular to B_0 (z – axis).

After the 90° pulse the nuclear magnetization vector will be rotating in the plane perpendicular to the magnetic field (xy plane). The rotating magnetic field will induce an electric current, which can be measured in circuit, the receiver coil, which is placed around the sample in the magnet.

1.3.1.4 *Energy levels, populations & signal-to-noise*

After inserting a NMR sample into the static magnetic field B_0 the spins in the ensemble of molecules in our NMR tube are either in the α - or β -state. That is the spins precess around the axis of the static magnetic field with a z -axis projection being parallel or antiparallel to B_0 respectively. The population of the α - and β -states according to a Boltzmann distribution is almost equal ($\sim 10^{-4}$ difference) since the energies involved are fairly small. Nevertheless the small population difference produces an effective

magnetization along the z-axis. Note, however, that there is no magnetization observed in the x,y-plane since the individual precession frequencies are uncorrelated, e.g. out of phase such that the x- or y- components average to zero. In order to measure the precession frequencies the system is disturbed and brought into a non-equilibrium state which allows to monitor the transverse (x or y) magnetization (Figure 1.7).

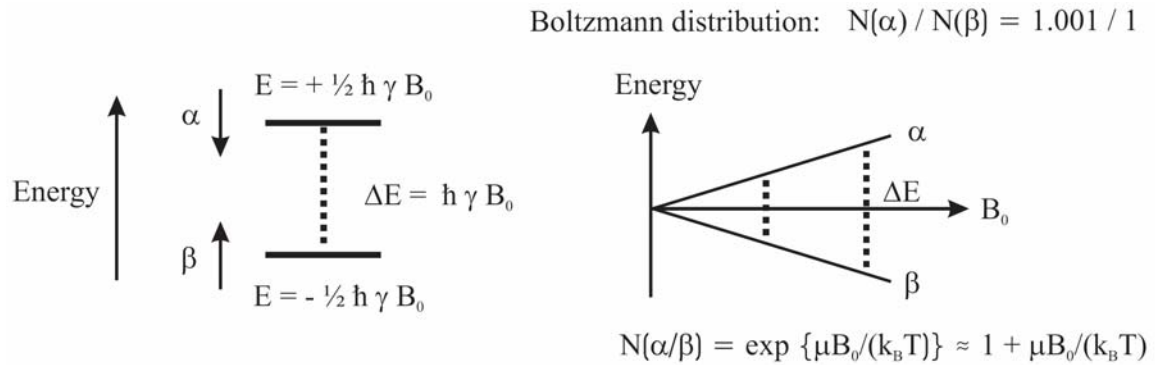


Figure 1.7. Energy levels of a nucleus in an external magnetic field depend on the strength of the field. Distribution between two energy states obeys Boltzmann law.

The Electromagnetic Spectrum

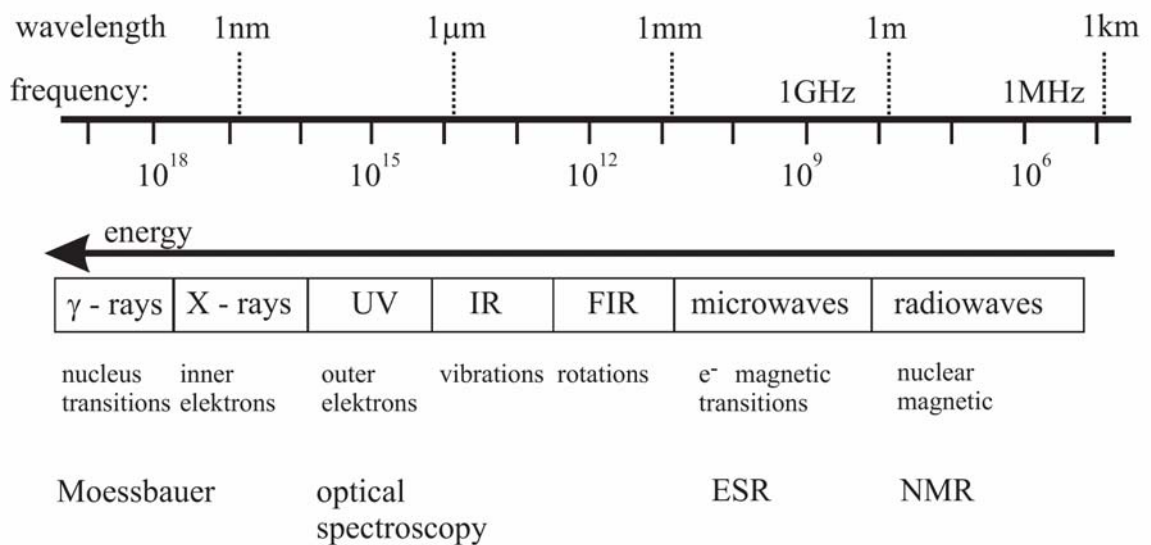


Figure 1.8. Electromagnetic wavelength spectrum.

NMR involves resonance frequencies that correspond to radiofrequency wavelengths and is thus located at the low frequency end of the electromagnetic spectrum. Since the separation of the energy levels (ΔE) is very small NMR is a rather insensitive spectroscopic method, and optimizing the signal-to-noise ratio is always a critical issue for NMR. The signal-to-noise (S/N) can be increased by using more spins (higher sample concentration), but also by using higher strengths for the static magnetic field (B_0), and optimizing the design of NMR experiments, e.g. using spins with the highest γ (such ^1H) for excitation and detection.

1.3.1.5 Nuclear magnetic resonance

Since in a protein the frequencies which correspond to the separation (E) of the spin energy levels vary, a number of frequencies have to be applied. There are two principal ways to achieve this. Either, the radiofrequency is varied within the expected frequency range, in the same way as a radiostation is tuned in (continuous wave, CW NMR), or a very short radiofrequency pulse is applied which inherently encodes a range of frequencies allowing to induce resonances for the whole frequency spectrum in one experiment (Fourier transform, FT NMR) (Figure 1.9). Since it takes time to step-wise tune over a range of frequencies by CW NMR while by FT NMR all signals can be excited at once, FT NMR is by far, more sensitive. The response obtained from a FT NMR experiment is a superposition of the frequencies of all spins in the molecule as a function of time, $F(t)$. In order to obtain the corresponding spectrum $F(\omega)$ (intensity as a function of frequency) a Fourier transformation is performed (Iannone 1999; Van Bramer 1999).

$$\begin{array}{ll} F(t) & \rightarrow F(\omega): & F(\omega) = \int F(t) \exp^{-i\omega t} dt \\ F(\omega) & \rightarrow F(t): & F(t) = \int F(\omega) \exp^{i\omega t} d\omega \end{array}$$

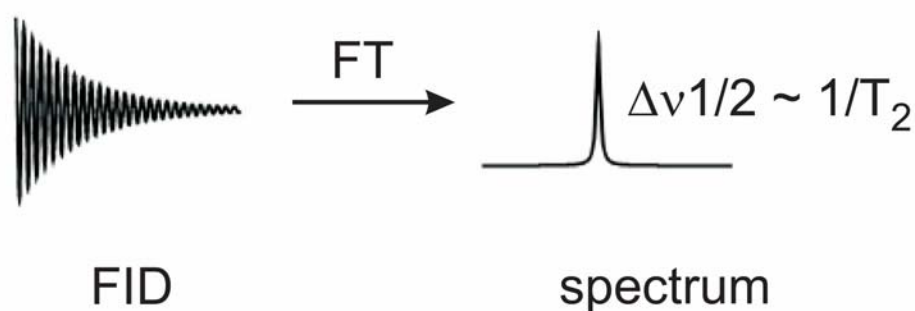


Figure 1.9. Fourier transformation.

1.3.1.6 Relaxation and NMR linewidths

After a 90° pulse has been applied to the equilibrium z-magnetization, transverse magnetization is created in which the phases of the individual spin precession frequencies are correlated. Such a nonrandom superposition of states is called coherence. The loss of phase-coherence in the transverse plane due to spin-spin interactions defines the T_2 or transverse relaxation time. In addition, the population difference along Z is restored by interactions with the surrounding lattice. This spin-lattice relaxation is described by the T_1 or longitudinal relaxation time. The linewidth of an NMR signal is described by the exponential decay of the FID, and reciprocal to the transverse relaxation time T_2 . It scales with the molecular tumbling rate in solution and therefore increases with higher molecular weight. This is the main reason for the molecular weight limitation of NMR.

1.3.1.7 Chemical Shifts

The magnetic field B_0 is shielded by the electrons in the local environment of a spin. Therefore, the individual resonance frequencies are slightly different reflecting the different chemical surrounding. The resonance frequencies are called chemical shifts and are measured in parts per million (ppm) in order to have chemical shift values independent of the static magnetic field strength (Figure 1.10). For example, backbone amide protons HN in a protein resonate around 8 ppm, while H_α spins have resonance frequencies between 3.5-5.5 ppm.

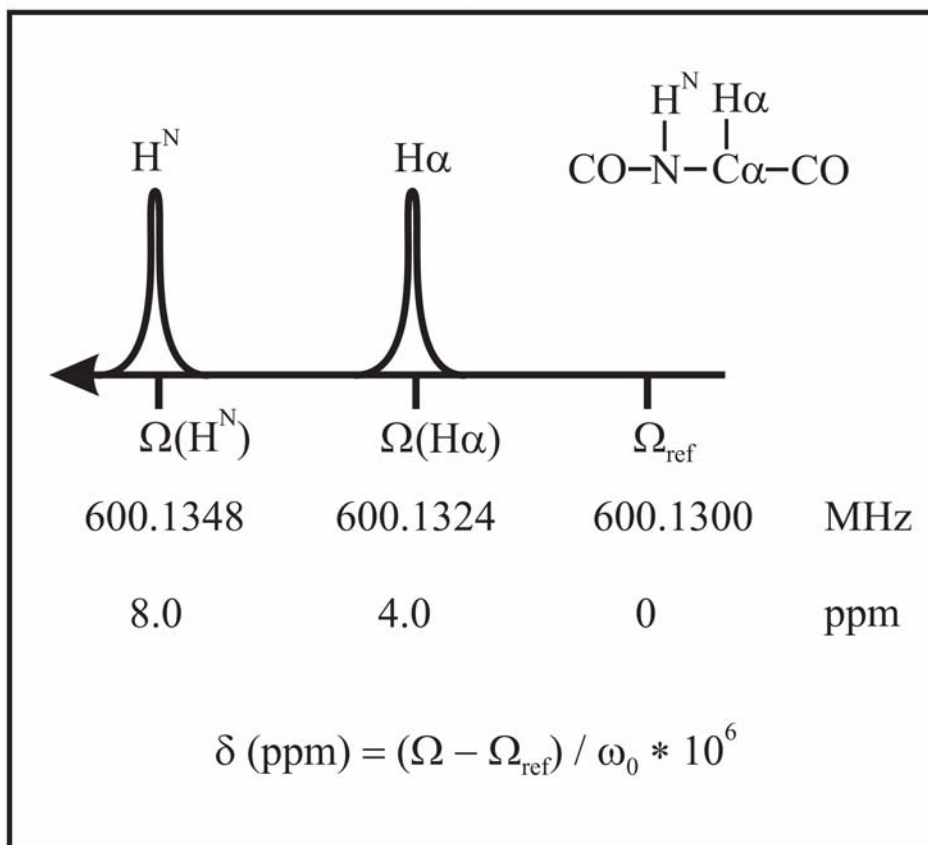


Figure 1.10. Chemical shift of H^{N} and H^{α} in ppm and corresponding values in MHz for the 600 MHz spectrometer.

1.3.1.8 J-coupling

Scalar or J-couplings are mediated through *chemical bonds* connecting two spins. The energy levels of each spin are slightly altered depending on the spin state of a scalar coupled spin (α or β) (Figure 1.11). This gives rise to a splitting of the resonance lines. Typical coupling constants in peptides are: ${}^1\text{J}(\text{H}^{\text{N}}, \text{N}) \sim 92$ Hz, ${}^3\text{J}(\text{H}^{\text{N}}, \text{H}^{\alpha}) \sim 2\text{-}10$ Hz, ${}^3\text{J}(\text{N}, \text{H}^{\beta}) 0\text{-}7$ Hz. ${}^3\text{J}$ couplings are well-correlated with the central dihedral angle by an empirical correlation, the Karplus curve. For example ${}^3\text{J}(\text{H}^{\text{N}}, \text{H}^{\alpha})$ defines the backbone angle ϕ in proteins. Scalar couplings are used in multidimensional (2D, 3D, 4D) correlation experiments to transfer magnetization from one spin to another in order to identify spin systems, e.g. spins which are connected by not more than three chemical bonds. For example, the amino acid ${}^1\text{H}$ spin system of alanine consists of H^{N} , H^{α} and H^{β} .

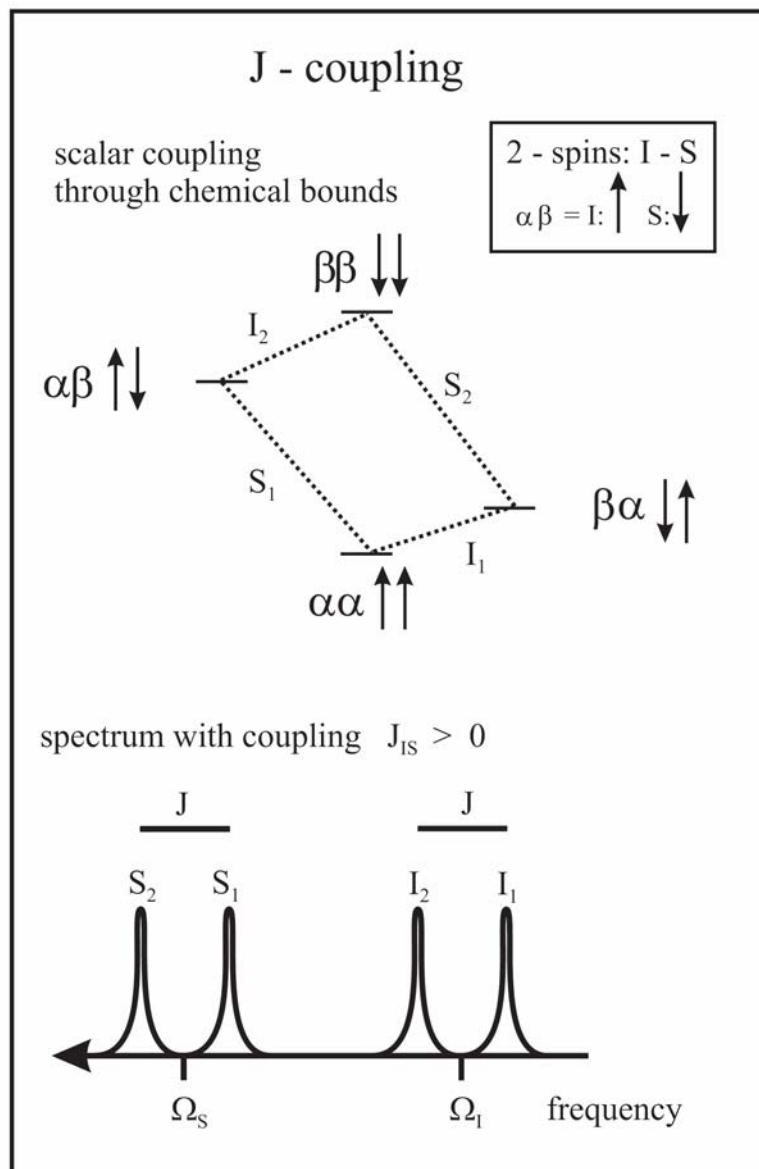


Figure 1.11. The 3J -coupling between two nucleus cause splitting of the signals in the NMR spectrum.

1.3.1.9 Nuclear Overhauser effect (NOE)

The nuclear Overhauser effect arises from cross relaxation between dipolar coupled spins as a result of spin/spin interactions through space. Dipolar couplings are usually in the kHz range, and depend on the distance between the two spins and the orientation of the internuclear vector with respect to the static magnetic field B_0 . Fortunately, due to the fast overall tumbling of molecules in solution, the dipolar couplings are averaged to zero. Nevertheless, the dipolar couplings give rise to spin/spin and spin/lattice relaxation. The NOE scales with the distance “r” between the two spins as $\sim 1/r^6$. Therefore the NOE is

related to the three-dimensional structure of a molecule. For interproton distances bigger than 5 Å, the NOE is too small and not observable.

1.3.1.10 Exchange

The exchange between two conformations, e.g. free and ligand bound forms of a protein, but also chemical exchange usually gives rise to two distinct NMR signals for a given spin due to different chemical environments in the two exchanging forms. If the exchange rate k is slow on the chemical shift time scale, two sets of signals are observed, if the exchange rate is fast on the chemical shift time scale only one signal is observed at an average frequency corresponding to the populations of the two conformations. If $k \sim \Delta\Omega$ intermediate exchange gives rise to very large line widths. From line shape analysis the exchange rate k and the populations of the conformations can be obtained. Since the frequency difference in Hz between the two signals is proportional to the magnetic field strength B_0 , the exchange regime may change between different B_0 field strengths.

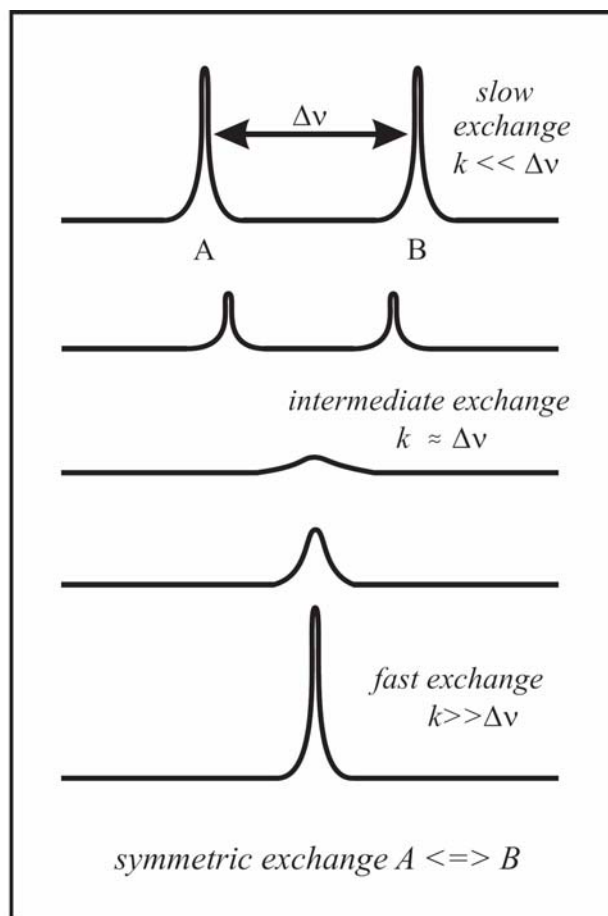


Figure 1.12. Exchange rate could be deduced from the NMR spectrum.

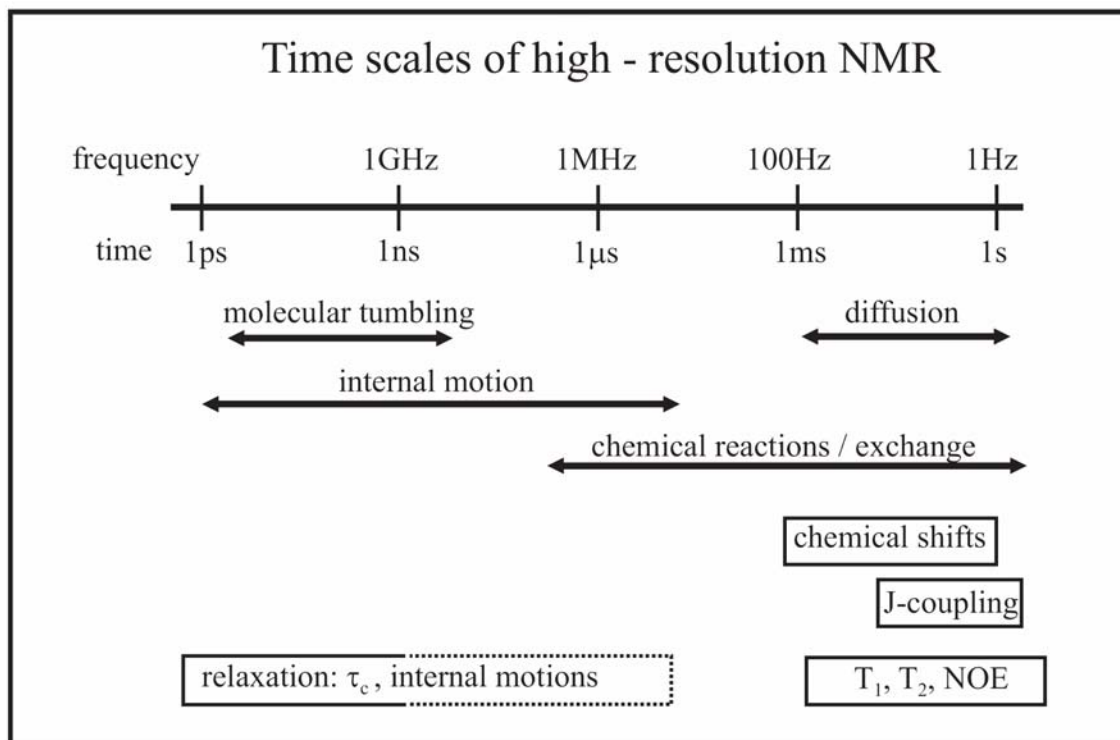
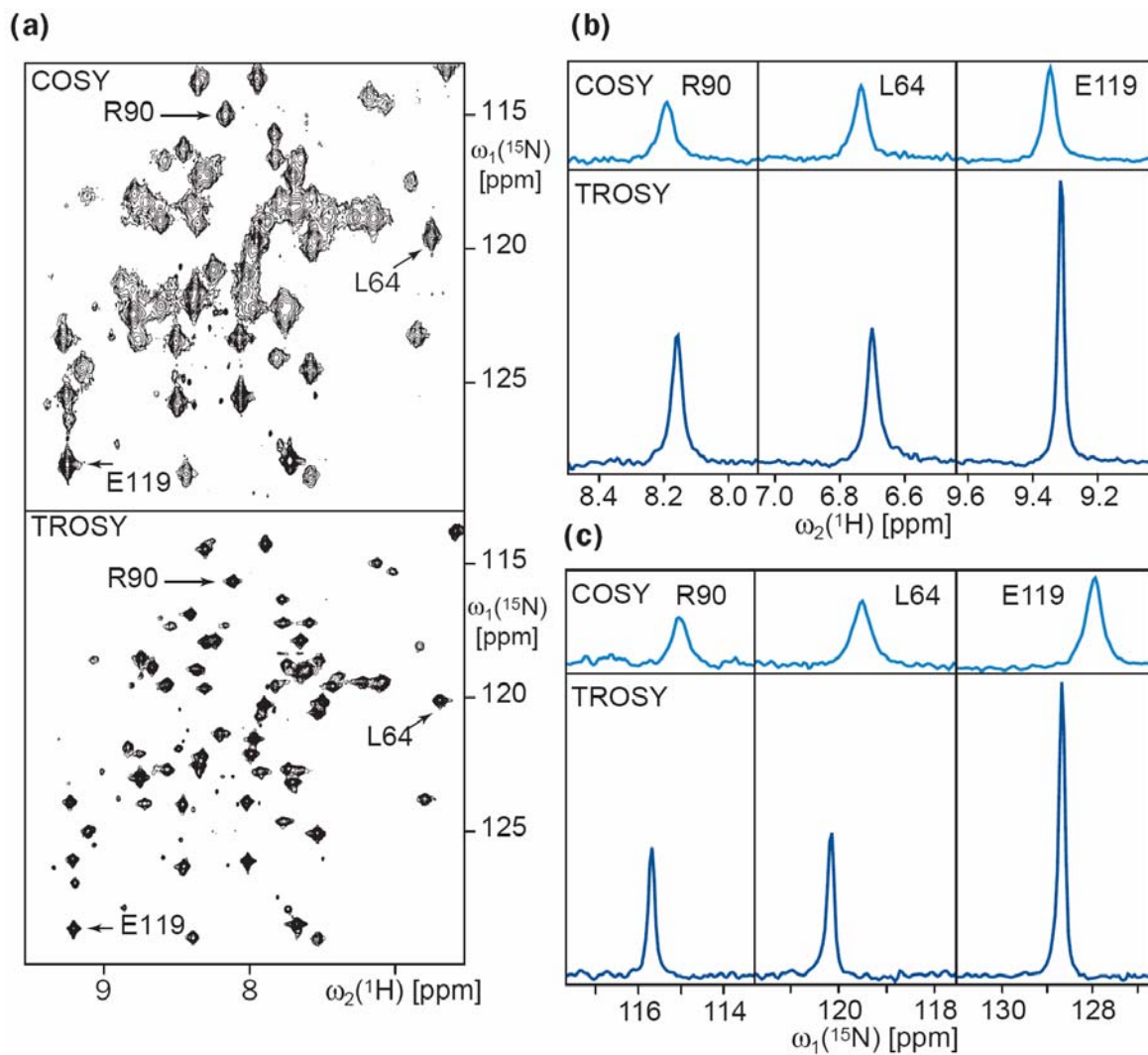


Figure 1.13. Time scale of NMR experiments and molecular processes.

1.3.1.11 Larger molecules, relaxation and TROSY

The main problem in studies of biomolecules with molecular weights above 30 kDa is the fast decay of the NMR signal due to relaxation. Therefore the signal-to-noise in NMR spectra of larger molecules is poor. The line width problem can be overcome (i) by the use of ^2H -labeling to eliminate ^1H mediated relaxation pathways and (ii) by a new NMR method called Transverse Relaxation Optimized Spectroscopy (TROSY) (Czisch et al. 1998; Pervushin et al. 2001; Fernandez et al. 2003). By the combined use of ^2H -labeling and TROSY, high-resolution NMR spectra of structures with molecular weights >100 kDa can be recorded. This, significantly extends the range of macromolecular systems that can be studied by NMR in solution. For example, spectra of molecules in the 900 kDa range have been reported (Fiaux et al. 2002). TROSY has already been used to map protein/protein interfaces, to conduct structural studies on membrane proteins and to study nucleic acid conformations in multimolecular assemblies. Thus, the conformational states of individual macromolecular components can be studied and *de novo* protein structure determination in large supramolecular structures will be feasible in the future.



Observable	Information
Chemical shifts ^1H , ^{13}C , ^{15}N , ^{31}P	Assignments, secondary structure (Wüthrich 1994)
J – couplings (through bond) $^3\text{J}(\text{H}^{\text{N}}, \text{H}\alpha)$, $^3\text{J}(\text{H}\alpha, \text{H}\beta)$, ...	Through-bond correlation experiments Dihedral angles: Φ , χ , Karplus curves (Wüthrich 1994).
NOE (through space)	Interatomic distances ($<5\text{\AA}$) (Neuhaus 2000)
Solvent exchange (H^{N})	Hydrogen bonds
Relaxation/linewidths ^1H , ^{13}C , ^{15}N cross-correlated relaxation (CCR)	Mobility, dynamics, conformational/chemical exchange, projection angles (Reif et al. 1997)
Residual dipolar couplings (RDCs) ^1H - ^{15}N , ^1H - ^{13}C , ^{13}C - ^{13}C , ...	Projection angles (Tjandra et al. 1997)

Table 1.4. Summary of the type of information one can obtain from different NMR experiments.

1.3.1.12 The NMR spectrum of a folded protein

In a protein structure the individual residues are packed into chemical environments, which are very different from the random coil situation. Figure 1.15 shows the NMR spectrum of a folded protein in comparison with the NMR spectrum of the same protein recorded at conditions where the protein is unfolded. The spectrum of the unfolded protein corresponds to a spectrum, which in essence is the sum of the random coil spectra of the amino acid residues in the proteins. The dispersion of signals in the spectrum of the folded protein is far beyond the envelope of signals seen in the spectrum of the unfolded protein. This clearly reflects that nuclei in the folded form are subject to a many different types of microenvironments of chemical screens.

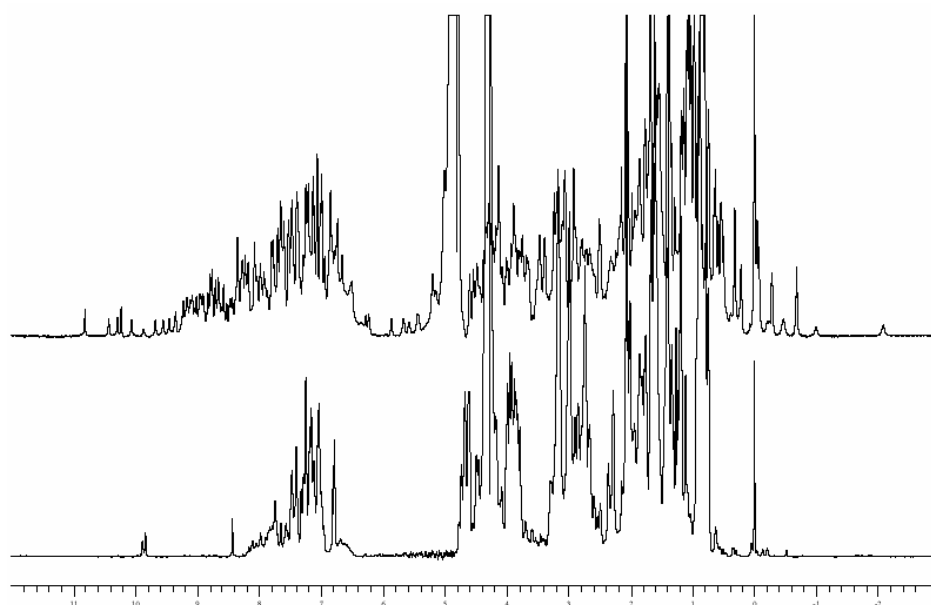


Figure 1.15. Comparison of NMR spectra of folded (top) and partially folded (bottom) protein.

Protons coupled to amine groups, aliphatic, aromatic, methyl or C α carbons are situated at the different position on the protein proton spectra (Figure 1.16).

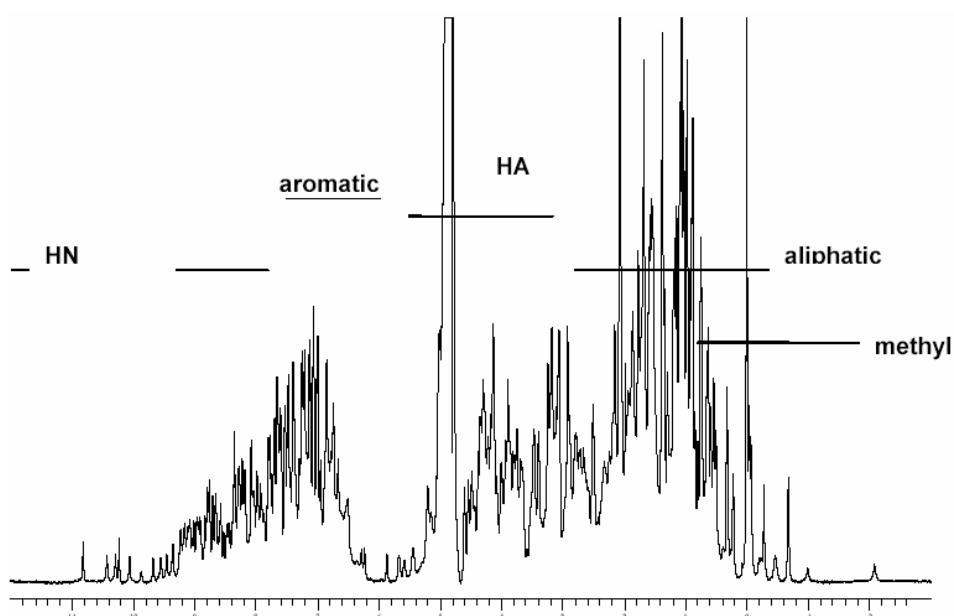


Figure 1.16. ^1H -NMR spectrum of hen egg white lysozyme.

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 *Chemicals*

All chemicals were of analytical grade, obtained from Merck (Germany) or Sigma Laboratories (USA), unless otherwise mentioned.

2.1.2 *Enzymes and buffers*

BSA 100x Stock, New England BioLabs (USA);

Pfu Turbo DNA Polymerase, 10x Cloned Pfu Reaction Buffer, Stratagene (USA);

Dpn I; Stratagene (USA);

DAPzyme, Qiagen (USA);

Xa factor, Sigma (USA);

Composition of other solutions and buffers used are described Sambrook (2001)

2.1.3 *Kits and reagents*

BD Advantage 2 PCR Kit, BD Biosciences (USA);

QIAquick PCR Purification Kit, Qiagen (Germany);

QIAprep Spin Miniprep Kit, Qiagen (Germany);

Rapid DNA Ligation Kit, Roche Applied Science (Germany);

Quick – Change Mutagenesis Kit, Stratagene (USA);

RTS-500, ROCHE (Germany)

2.1.4 *Oligonucleotides*

Oligonucleotides used in cloning were provided by Metabion, or MWG (Germany). The primers are listed below.

2.1.4.1 *Primers for MyoD mutagenesis*

1. **s_C135S+C7**

5' - AGGCTCTGCTGCGCGACCAGGACGCCGCCCGGGCTAAGGCTCTAACTCT - 3'

2. **as_C135S+C7**

5' - AGAGTTAGAGCCTTAGCCCCGGCGGGGCGGCGTCCCTGGTTCGCGCAGCAGAGCCT - 3'

3. **s_C135S+N4**

5' - ATCATCATCATCGCAAAGCCTGCAAACGCAAGACCACCAACGCCGA - 3'

4. **as_C135S+N4**

5' - TCGGCGTTGGTGGTCTTGCCTTTGCAGGCTTTGCGATGATGATGAT - 3'

5. **MyoD_s_S135C**

5' -GAGACACTCAAGCGCAGCACGTCGAGCAATCCA-3'

6. **MyoD_as_S135C**

5' -TGGATTGCTCGACGTGCTGCGCTTGAGTGTCTC-3'

7. **MyoD_s_del_c**

5' - GGAGATATACATATGAAGCGCAAGACCACCAACG - 3'

8. **MyoD_as_del_c**

5' - CGTTGGTGGTCTTGCCTTCATATGTATATCTCC - 3'

9. **MyoD_s_C135S**

5' -GAGACACTCAAGCGCTGCACGTCGAGCAATCCA - 3'

10. **MyoD_as_C135S**

5' -TGGATTGCTCGACGTGCAGCGCTTGAGTGTCTC-3'

*2.1.4.2 Primers for MyoD cloning to different vectors**2.1.4.2.1 pET30 Xa/LIC*1. **MyoD_as**

5' - AGAGGAGAGTTAGAGCCTTAGTCGCGCAGCAGAGCCTGCAGGCCCTCGAT - 3'

2. **MyoD_s**

5' - GGTATTGAGGGTCGCAAGACCACCAACGCCGACCGCCGC - 3'

*2.1.4.2.2 pET41 ek/LIC*1. **MGSTlig_as**

5' - GAGGAGAAGCCCCGTTAGAGCACCTGGTATATCGGGTTGGGGTTTCG - 3'

2. **MGSTlig_s**

5' - GACGACGACAAGATGGAGCTACTGTCGCCACCGCTC - 3'

*2.1.4.3 Primers for MyoD gen synthesis*1. **MyoD_C17_as**

5' - AGAGGAGAGTTAGAGCCTTAGCCTTCGTAGCAGTTGCGGCGGCG

GGCGCCGCTCGGCGGGCCGCTGTAGTCCATGCG - 3'

2. **MyoD_C17_s**

5' - GGTATTGAGGGTTCGCATGGACTACAGCGGCCCGCCGAGCGGCGC
CCGCCGCCGCAACTGCTACGAAGGCTAAGG -3'

2.1.4.4 *Primers for Id – 2 cloning*

2.1.4.4.1 *pET30 Xa/LIC*

1. **Id2lig_as**

5' - AGAGGAGAGTTAGAGCCTTACAGTGCTTTGCTGTCATTTGACATTAAGTC -3'

2. **Id2lig_s**

5' - GGTATTGAGGGTTCGCATGAAAGCCTTCAGTCCCGTGAGG -3'

2.1.4.4.2 *pET41 ek/LIC*

1. **ID2GSTLIG_AS**

5' - GAGGAGAAGCCCGTTACAGTGCTTTGCTGTCATTTGACATTAAGTC -3'

2. **ID2GSTlig_s**

5' - GACGACGACAAGATGAAAGCCTTCAGTCCCGTGAGG -3'

2.1.4.5 *Primers for Id – 2 mutagenesis*

1. **ID2_as_C52S**

5' - GCTCCTTGAGCTTGGAGTAGGAGTCGTTTCATGTTGTATAGC - 3'

2. **ID2_s_C52S**

5' - GCTATACAACATGAACGACTCCTACTCCAAGCTCAAGGAGC - 3'

3. **s_ID2Cys**

5' - GCTATACAACATGAACGACTGCTACTCCAAGCTCAAGGAGC - 3'

4. **as_ID2Cys**

5' - GCTCCTTGAGCTTGGAGTAGCAGTCGTTTCATGTTGTATAGC - 3'

5. **ID2_s_Cdel**

5' - CTGCAGATCGCCCTGGACTCGTAAGGCTCTAACTCTCCTCT - 3'

6. **ID2_as_Cdel**

5' - AGAGGAGAGTTAGAGCCTTACGAGTCCAGGGCGATCTGCAG -3'

7. **ID2_s_Ndel**

5' - TGCACCATCATCATCATCATAAGAAAAGCCTGGGCATCTCCCGGAGC -3'

8. **ID2_as_Ndel**

5' - GCTCCGGGAGATGCCAGGCTTTTCTTATGATGATGATGATGGTGCA -3'

9. **s_ID2_N2B2_N-6**
5' - TGCACCATCATCATCATCATAAGCGGAGCAAAACCCCTGTGGAC -3'
10. **as_ID2_N2B2_N-6**
5' - GTCCACAGGGGTTTTGCTCCGCTTATGATGATGATGATGGTGCA -3'
11. **s_ID2_N2B2_C-45**
5' - ACTCGCATCCCCTATTGTCAGCTAACTGCATCACCAGAGACC -3'
12. **as_ID2_N2B2_C-45**
5' - GGTCTCTGGTGATGCAGTTAGCTGACAATAGTGGGATGCGAGT -3'
13. **s_ID2_N2B2_C-35t**
5' - GGCAGAACCAGTAAGCGTCCAGGACGCCGCTGA -3'
14. **as_ID2_N2B2_C-35t**
5' - TCAGCGGCGTCCTGGACGCTTACTGGTTCTGCC -3'
15. **s_ID2_N2B2_C-35d**
5' - GGCAGAACCAGTAAGGCTCTAACTCTCCTCTG -3'
16. **as_ID2_N2B2_C-35d**
5' - CAGAGGAGAGTTAGAGCCTTACTGGTTCTGCC -3'
17. **s_ID2_N2B2_C-25**
5' - GCTGACCACCCTCAACTAAACGGATATCAGCATCCTG -3'
18. **as_ID2_N2B2_C-25**
5' - CAGGATGCTGATATCCGTTTAGTTGAGGGTGGTCAGC -3'

2.1.4.6 Primers for E7 gene synthesis

1. **s_E71_pET30**
5' - GGTATTGAGGGTTCGCATGGATCTGTATTGTTATGAACAGCTGAACGATTCTTCTGAAGAA
GAAGATGAAATTGATGGTTAAGG -3'
2. **as_E71_pET30**
5' - AGAGGAGAGTTAGAGCCTTAACCATCAATTTTCATCTTCTTCTTTCAGAAGAATCGTTCAGCT
GTTTCATAACAATACAGATCCATGCG -3'
3. **s_E72_pET30**
5' - GGTATTGAGGGTTCGCATGGATCTGTATTGTTATGAACAGCTGAACGATTCTTCTGAAGAAG
AAGATGAAATTGATGGTCCGGCGGGTTCAGGCTGAACCGGATCGTGCTTAAGG -3'
4. **as_E72_pET30**
5' - AGAGGAGAGTTAGAGCCTTAGCACGATCCGGTTCAGCCTGACCCGCCGACCATCAATTTCA
TCTTCTTCTTTCAGAAGAATCGTTCAGCTGTTTCATAACAATACAGATCCATGCG -3'

5. **s_E71_pET41**

5' - GACGACGACAAGATGGATCTGTATTGTTATGAACAGCTGAACGATTCTTCTGAAGAAGAAGATGAAATTGATGGTTAA -3'

6. **as_E71_pET41**

5' - GAGGAGAAGCCCGTTAACCATCAATTTTCATCTTCTTCTTCAGAAGAATCGTTCAGCTGTTCATAACAATACAGATCCA -3'

7. **s_E72_pET41**

5' - GACGACGACAAGATGGATCTGTATTGTTATGAACAGCTGAACGATTCTTCTGAAGAAGAAGATGAAATTGATGGT -3'

8. **as_E72_pET41**

5' - GAGGAGAAGCCCGTTAGCACGATCCGGTTCAGCCTGACCCGCCGGACCATCAATTTTCATCTTCTTCTTCAGAAGAATCGTTCAGCTGTTCATAACAATACAGATCCA -3'

2.1.4.7 *Primers for E7 mutagenesis*1. **s_E73_del**

5' - ATGCACGGTGACACCGCTGGTCAGGCTGAACCG -3'

2. **as_E73_del**

5' - CGGTTTCAGCCTGACCAGCGGTGTCACCGTGCAT -3'

2.1.4.8 *Primers for GFP mutagenesis*2. **s_Y66W**

5' - GCCAACACTACTTGTCACTACTTTCTCTTGGGGTGTTC AATGCTTTTTCCCG -3'

3. **as_Y66W**

5' - CGGGAAAAGCATTGAACACCCCAAGAGAAAAGTAGTGACAAGTGTGGC -3'

4. **s_Y66H**

5' - GCCAACACTACTTGTCACTACTTTCTCTCATGGTGTTC AATGCTTTTTCCCG -3'

5. **as_Y66H**

5' - CGGGAAAAGCATTGAACACCCATGAGAGAAAAGTAGTGACAAGTGTGGC -3'

2.1.5 *Plasmids and constructs*2.1.5.1 *Plasmids supplied by the companies*

pET30LIC/Xa (Novagen);

pET41LIC/Ent (Novagen);

pIVEX (ROCHE);

pET3 (Novagen);
pET8 (Novagen)

2.1.5.2 Plasmids constructed during the work

Most of the plasmids used for this work were based on pET30LIC/Xa or pET41LIC/Ent. Both have N-terminal tags. The first has a 6 histidine tag (His-tag), the second has a GST tag.

All Id-2, MyoD and most of pRb constructs were cloned into pET30 Xa/LIC vector (NOVAGEN) according to manufacturer manual. Short His-tag and DAPzyme digestion stop site were introduced by Quick-Change (Stratagen) according to manufacturer manual. CDK6 construct was cloned between BamHI and EcoRI sites into pAcG2T vector. MCNC construct consisted of 6 His, DAPzyme digestion stop site and MyoD sequence 99–173. pRb ABC-C consisted of 6 His, DAPzyme digestion stop site and pRb sequence 379–791. pRb ABC-CL construct spanned residues 379–578 and 642–791 of pRb. CDK6 construct consisted of full length CDK6 cloned on the C-terminal side of GST. MyoD18C construct spanned residues 211–227 of MyoD and MyoD25C spanned residues 201–232. Id-2 construct covers residues 2–132. HPV16 E7 full length protein was a gift from Roland Degenkolbe. The p19 construct spanned the full length of the protein and the E2F residues 243–437.

The plasmid pET8 containing full length pRb was a kind gift from Prof. Wen-Hwa Lee (University of Texas Health Science Center, USA). The plasmid p2.4, containing A, B and C domain without linker between first two domains and a stop codon introduced after B domain, was prepared in laboratory.

GFP constructs were cloned to pET3 vector. They consist of full length GFPuv (synonymous for GFP cycle3) with six histidine residues at the C-terminus of protein.

Other constructs are mentioned in the tables from 3.1 – 3.5.

2.1.5.2.1 *Constructs of MyoD protein*

Name	Tag	Tag length	Tag seq	N-term	Ser/Cys	C-term	Length
MyoD	HIS	9	MHHHHHHRK	102	Ser	163	61
MCN	HIS	8	MHHHHHHR	99	Ser	173	74
MCNC	HIS	8	MHHHHHHR	99	Cys	173	74
MyoDC	HIS	9	MHHHHHHRK	102	Ser	173	71
MyoDN	HIS	8	MHHHHHHR	99	Ser	163	64
GMyoD	GST			1	Cys	319	318

Table 2.1. Constructs of MyoD.

2.1.5.2.2 *Constructs of ID2 protein*

Name	Tag	Tag length	Tag seq	N-term	Ser/Cys	C-term	Length
N12	HIS	8	MHHHHHHK	19	Ser	81	62
R2BS	HIS	8	MHHHHHHK	2	Ser	132	130
R2B	HIS	8	MHHHHHHK	2	Cys	132	130
N2B2	HIS	9	MHHHHHHKK	19	Ser	132	113
IC25	HIS	9	MHHHHHHKK	19	Ser	107	88
IC35	HIS	9	MHHHHHHKK	19	Ser	97	78
IC45	HIS	9	MHHHHHHKK	19	Ser	87	68
IN4	HIS	9	MHHHHHHKK	23	Ser	132	109
GST-ID2	GST			1	Ser	134	134
GID2-C	GST			1	Cys	134	134
IC45C	HIS	9	MHHHHHHKK	19	Cys	87	68
R2BS	HIS	8	MHHHHHHK	2	Ser	132	130

Table 2.2. Constructs of Id-2 protein.

2.1.5.2.3 Constructs of E7 protein

Name	Tag	N-term	C-term	Length
E713	HIS	21	40	20
E723	HIS	21	50	30
E714	GST	21	40	20
E724	GST	21	50	30
E7	HIS	1	98	98

Table 2.3. Constructs of Id-2 protein.

2.1.5.2.4 Plasmid and various construct GFP for Protein Expression

Construct	Description
GFPuv	pET3 expressing vector with GFPuv (also called GFP cycle3) sequence followed by C – terminal HIS – tag.
GGFPuv	pET3 expressing vector with GFPuv (also called GFP cycle3) sequence followed by C – terminal HIS – tag. Sequence is mutated to give H148G substitution.
HGFP	pET3 expressing vector with GFPuv (also called GFP cycle3) sequence followed by C – terminal HIS – tag. Sequence is mutated to give Y66H substitution.
WGFP	pET3 expressing vector with GFPuv (also called GFP cycle3) sequence followed by C – terminal HIS – tag. Sequence is mutated to give Y66W substitution.

Table 2.4. Constructs of GFP protein.

2.1.5.2.5 Constructs of pRb protein

pET30LIC/Xa-ABC	RB ABC domain with the spacer region between A & B (from AA 379 to 928)
pET30LIC/Xa-ABC-C	RB, AB domain with the spacer region between A & B (from AA 379 to 792)
pET30LIC/Xa-ABC-CL	RB A & B domain without the spacer region between A & B (379–578 and 642–791)

Table 2.5. Constructs of pRb protein.

2.1.5.2.6 Constructs of α – PEC

The SamI-XhoI fragment of α – PEC was cloned to expression vector pET-30(+) (Storf et al. 2001).

2.1.6 Bacterial strains

The strains are listed in Table 2.6. Cultures were grown in LB medium (Sigma, USA). Whenever needed, media were supplemented with kanamycin, ampicillin or chloramphenicol in working concentrations of 150 $\mu\text{g/ml}$, 100 $\mu\text{g/ml}$ and 34 $\mu\text{g/ml}$, respectively, unless otherwise mentioned.

Bacterial strains	Genotype
TOP10	F ⁻ <i>mcrA</i> $\Delta(mrr\ hsdRMS\ mcrBC)$ $\Phi80lacZ\Delta M15$ $\Delta lacX74$ <i>deoR</i> <i>recA1</i> <i>araD139</i> $\Delta(ara-leu)7697$ <i>galU</i> <i>galK</i> <i>rpsL</i> (Str ^R) <i>endA1</i> <i>nupG</i>
BL21 Star(DE3)	F ⁻ <i>ompT</i> <i>hsdS_B</i> (<i>r_B⁻</i> , <i>m_B⁻</i>) <i>gal dcm me131</i> (DE3)
BL21 Star(DE3)pLysS	F ⁻ <i>ompT</i> <i>hsdS_B</i> (<i>r_B⁻</i> , <i>m_B⁻</i>) <i>gal dcm me131</i> (DE3) pLysS (Cam ^R)
BL21 Star(DE3)pLysE	F ⁻ <i>ompT</i> <i>hsdS_B</i> (<i>r_B⁻</i> , <i>m_B⁻</i>) <i>gal dcm me131</i> (DE3) pLysE (Cam ^R)
Rosetta(DE3)	F ⁻ <i>ompT</i> <i>hsdS_B</i> (<i>r_B⁻</i> , <i>m_B⁻</i>) <i>gal dcm lacY1</i> (DE3) pRARE(Cam ^R)
BL21 CodonPlus (DE3)-RP	F ⁻ <i>ompT</i> <i>hsdS_B</i> (<i>r_B⁻</i> , <i>m_B⁻</i>) <i>gal dcm Tet^R</i> (DE3) <i>endA Hte</i> [<i>argU proL</i> Cam ^R]
BL21 Gold(DE3)pLysS	F ⁻ <i>ompT</i> <i>hsdS_B</i> (<i>r_B⁻</i> , <i>m_B⁻</i>) <i>gal dcm Tet^R</i> (DE3) <i>endA Hte</i> pLysS (Cam ^R)

Table 2.6. Bacterial strains used in this work.

2.1.7 Peptides

Peptides were synthesised using solid phase, purified with C8 reverse phase chromatography and checked by mass-spectroscopy.

Peptide name	Corresponding protein	Sequence
h – MyoD - C	human MyoD (212–227)	YSGPPSGARRRNCYEG
ch – MyoD - C	chicken MyoD (189–204)	YSGPPCSSRRRNSYDS
E7-0	HPV16 E7 (21–29)	DLYCYEQLN
E7-1	HPV16 E7 (21–40)	DLYCYEQLNDSSEEEDEIG
Large – T	SV40 large T antigen	NLFCSEEMD

Table 2.7. Peptides corresponding to fragments of different proteins were synthesized.

2.1.8 *Buffers and media*

All buffers, stock solutions and media, if not mentioned here, were performed exactly like described in Sambrook & Russell (2001).

2.1.8.1 *LB medium:*

Tryptone 10g/l

Yeast Extract 5g/l

NaCl 5g/l

Adjust the pH to 7-7.5

For the preparation of agar plates the medium was supplemented with 15g agar. Antibiotic were added after the medium has been cooled to 50°C.

2.1.8.2 *Minimal medium (MM) for uniform enrichment with ¹⁵N*

Stock solutions:

1. thiamin 1%
2. MgSO₄ 1 M
3. Zn-EDTA solution:
 - a. EDTA 5 mg/ml
 - b. Zn(Ac)₂ 8.4 mg/ml

Dissolved separately in small water volumes, and then mixed together.

4. trace elements solution:
 - a. H₃BO₃ 2.5 g/l
 - b. CoCl₂*H₂O 2.0 g/l
 - c. CuCl₂*H₂O 1.13 g/l
 - d. MnCl₂*2H₂O 9.8 g/l
 - e. Na₂MoO₄*2H₂O 2.0 g/l

If difficult to dissolve, pH was lowered with citric acid or HCl.

6. glucose 20%

Was separately autoclaved.

For 1 liter medium:

- | | |
|---|---|
| 1. NaCl | 0.5 g |
| 2. trace elements solution | 1.3 ml |
| 3. citric acid monohydrate | 1 g |
| 4. ferrous citrate | 36mg (dissolved in 120µl conc. HCl, heated) |
| 5. KH ₂ PO ₄ | 4.02 g |
| 6. K ₂ HPO ₄ *3H ₂ O | 7.82 g |
| 7. Zn-EDTA solution | 1 ml |
| 8. NH ₄ Cl or ¹⁵ NH ₄ Cl | 1 g |

pH was adjusted to 7.0 with NaOH

the mixture was autoclaved

25 ml separately autoclaved glucose was added

other compounds were added (previously sterile filtered):

- | | |
|---|--------|
| 1. thiamin solution | 560 µl |
| 2. antibiotics (half of the usual amount for LB-medium) | |
| 3. MgSO ₄ solution | 2 ml |

2.1.8.3 IPTG stock solution:

IPTG was dissolved in water (2.38 g / 10 ml) to the end concentration of 1 M. The stock solution was sterile filtered(0.22 µ) and stored in aliquots at -20°C until used. The final concentration of IPTG used in the culture was 1mM, unless otherwise indicated.

2.1.8.4 Antibiotics

Ampicilin, kanamycin, chloramphenicol stock solutions were made to the end concentrations of 100, 50, 50 mg / ml respectively. After filtration stored at -20 °C and used in 1000 dilutions in respective medium.

2.1.9 *Antibodies and other proteins*

2.1.9.1 *Antibodies*

Against	Name	Supplier
His-tag	H3-anti His probe	Santa Cruz Biotechnology, Inc., USA
pRb	IF8-anti pRb MAb	Santa Cruz Biotechnology, Inc., USA
Goat anti-mouse	Anti Goat	Santa Cruz Biotechnology, Inc., USA

Table 2.8. Antibodies used are listed.

2.1.9.2 *Molecular weight marker for SDS-PAGE electrophoresis*

Protein	Source	Apparent MW(Da)
Maltose-binding protein- β -galactosidase	E.Coli	175 000
Maltose-binding protein-paramyosin	E.Coli	83 000
Glutamic dehydrogenase	Bovine liver	62 000
Aldolase	Rabbit muscle	47 500
Triosephosphate isomerase	Rabbit muscle	32 500
γ -lactoglobulin A 25 000	Bovine milk	25 000
lysozyme	Chicken egg white	16 500
aprotinin Bovine lung 6 500	Bovine lung	6 500

Table 2.9. Protein used as a mass standard for SDS – PAGE (NEB, FRG).

2.1.10 *Other chemicals*

2.1.10.1 *Protease inhibitors:*

Complete Protease Inhibitors Cocktail (Roche, FRG)

2.1.10.2 Isotopically enriched chemicals:

¹⁵N-Ammonium chloride, NH₄Cl 99.9% (Campro Scientific, Berlin, FRG);

¹³C – Glucose (Campro Scientific, Berlin, FRG);

¹⁵N His, ¹⁵N Lys, ¹⁵N Leu (Campro Scientific, Berlin, FRG)

2.1.10.3 Miscellaneous :

Acetic acid; Acrylamide; L-Arginine; Ammonium chloride, NH₄Cl; Ammonium persulfate, APS; Biotin; Boric acid, H₃BO₃; Calcium chloride, CaCl₂; Citric acid; Cobalt (II) chloride, CoCl₂; Coomassie Brilliant Blue R-250; Copper (II) chloride, CuCl₂; Dimethylsulfoxide, DMSO; Dipotassium hydrogenphosphate, K₂HPO₄; Disodium hydrogenphosphate, Na₂HPO₄; Dithiothreitol, DTT; Ethanol; Ethylendiamintetraacetic acid, disodium salt, EDTA; Formaldehyde; Ferrous citrate; D-Glucose; Glutardialdehyde; L-Glutathione, oxidized, GSSG; L-Glutathione, reduced, GSH; Glycerine; Glycine; Guanidine hydrochloride; Hydrochloric acid, HCl; Imidazole; Isopropanol; Isopropyl-β-D-thiogalactopyranoside, IPTG; Magnesium chloride, MgCl₂; Magnesium sulfate, MgSO₄; Manganese (II) chloride, MnCl₂; β-Mercaptoethanol, β-ME; Methanol; N,N'-Methylenbisacrylamide; Nonidet P-40, NP-40; Potassium chloride, KCl; Potassium dihydrogenphosphate, KH₂PO₄; Silver nitrate, AgNO₃; Sodium acetate; Sodium azide, NaN₃; Sodium carbonate, Na₂CO₃; Sodium chloride, NaCl; Sodium dihydrogenphosphate, NaH₂PO₄; Sodium dodecylsulphate, SDS; Sodium hydrogencarbonate, NaHCO₃; Sodium hydroxide, NaOH; Sodium molybdate, Na₂MoO₄; Sodium thiosulfate; N,N,N',N'-Tetramethylenethyldiamine, TEMED; Thiamin; Tricine; Trifluoroethanol, TFE; Tris-(hydroxymethyl)-aminomethane, TRIS; Triton X-100; Tryptone; Yeast Extract; Zinc acetate, Zn(Ac)₂ (all Sigma-Aldrich Chemie, St Louis, USA)

2.2 Apparatus

2.2.1 *ÄKTA explorer 10 purification system*

ÄKTAexplorer 10 (Amersham Biosciences (Sweden)) is a complete fully automated liquid chromatography system consisting separation unit, fraction collector Frac-901 and a personal computer running UNICORN software v.3.1 to control the separation unit. The system enables to monitor: light absorption at 3 wavelengths simultaneously, in the range

190-700 nm; pH; conductivity; temperature; flow rate and pressure values. High performance pumps make possible to apply flow rates in range 0.01-10 ml and creating of gradients.

Samples were applied with 10 or 50 ml superloop (Amersham Biosciences, Sweden), which was filled manually. UV absorption was monitored at wavelengths 255, 270 and 280. As the pH of buffers was known and constant, pH was not monitored. Runs were performed in cold cabinet, at 4°C.

2.2.1.1 Chromatography equipment, columns and media:

NiNTA-agarose (QIAGEN, FRG);

ÄKTA explorer 10; Peristaltic pump P-1; Fraction collector RediFrac; ARecorder REC-1; UV flow through detector UV-1; HiLoad 16/60 Superdex S30pg, S200pg; HiLoad 26/60 Superdex S75pg; HiLoad 10/30 Superdex S75pg; Mono Q HR 5/5, 10/10; Mono S HR 5/5, 10/10; Buthyl Sepharose 4 FF; Q-Sepharose FF; SP-Sepharose FF; Glutathione Sepharose (all Amersham Pharmacia, Freiburg)

2.2.2 NMR spectrometers

600 MHz Bruker (Switzerland);

750 MHz Bruker (Switzerland);

600 MHz Ultrashielded Bruker (Switzerland);

900 MHz Ultrashielde Bruker (Switzerland)

2.2.3 Other apparatus

BioloLogic LP System Biorad, (FRG);

Centrifuges: J-6M/E and Avanti J-30I, Beckman (UK);

Incubator shaker: Lab-Therm, Kühner (Switzerland);

PCR machine: Mastercycler personal, Eppendorf (FRG);

Sonicator: Sonifer 250, Branson (USA);

Spectrophotometer: Ultraspec 3100 pro, Amersham Biosciences (Sweden);

LPLC system: peristaltic pump (Pump P-1), Single path monitor UV-1, recorder (REC-1), fraction collector (RediFrac), Amersham Biosciences (Sweden);

BIO-RAD MINI-PROTEAN II gel apparatus, Bio-Rad Laboratories (UK);

Autoclave Bachofer,	Reutlingen, (FRG);
Balances PE 1600, AE 163	Mettler, (FRG);
Centrifuge Avanti J-30I	Beckman, (USA);
Centrifuge Microfuge R	Beckman, (USA);
Centrifuge 3K15	Sigma, (FRG);
Centrifuge 5414	Eppendorf, (FRG);
Chambers for SDS PAGE and Western blotting	BioRad (FRG);
Ice machine Scotsman AF 30	Frimont, Bettolino di Pogliano, (Italy);
MARresearch image plates, mar345	MARresearch, Hamburg, (FRG);
Magnetic stirrer Heidolph M2000	Bachofer, Reutlingen, (FRG);
NMR-spectrometer DRX500, DRX600	Bruker, Rheinstetten, (FRG);
pH-meter pHM83	Radiometer, Copenhagen, (Denmark);
Pipettes 2.5µl, 10µl, 20µl, 200µl, 1000µl	Eppendorf, (FRG);
Quarz cuvettes	QS Hellma, (FRG);
Shaker	Adolf-Kühner AG, (Switzerland);
Spectrophotometer	Amersham Pharmacia, (FRG);
Ultrafiltration cells, 10ml, 50ml, 200ml	Amicon, Witten, (FRG);
Vortex	Cenco, (FRG);
Western blot apparatus (semidry)	MPI for biochemistry;
X-ray generator RU2000, 45kV, 120mA	Rigaku, Tokyo, (Japan)

2.3 Consumables

Centripreps YM3, YM10	Amicon, Witten, (FRG);
Dialysis tubing Spectra/Por MW 3500	Roth, Kleinfeld, Hannover, (FRG);
Falcon tubes, 15ml, 50ml	Becton Dickinson, Heidelberg, (FRG);
Maxi-Prep, Plasmid Isolation Kit	Qiagen, (FRG);
NMR-tubes, 5mm	Wilmad, Buena, NJ, (USA);
Parafilm	American National, (Canada);
Pipette tips 10µl, 200µl, 1000µl	Gilson, Villiers-le Bel, (France);
Plastic disposable pipettes 1ml, 5ml, 10ml, 25ml	Falcon, (FRG);
Reaction cups 0.4ml, 1.5ml, 2ml	Eppendorf, (FRG);
Sterile filters Millex 0.22µm, 0.45µm Millipore,	Molsheim, (FRG);

Syringes 1ml, 2ml, 10ml, 20ml, 60ml	Braun, Melsungen (FRG);
Ultrafiltration membranes YM3, YM10	Amicon, Witten, (FRG);

2.4 Methods

2.4.1 *DNA techniques*

2.4.1.1 *Basic cloning techniques*

All employed molecular biology protocols, if not mentioned here, were used exactly like described in Sambrook & Russell (2001).

2.4.1.2 *Isolation of the plasmids*

All plasmids were isolated using QIAprep Spin Miniprep Kit according to the protocol.

2.4.1.3 *Screening of positive colonies*

Single colony was picked up from the plate and transfer to 10 ml of LB media containing appropriate antibiotic and grown overnight at 37°C. From 1,5 ml of this culture plasmid was isolated as described above, and then either double digested with restriction enzymes or PCR was performed using isolated plasmid as the template.

2.4.1.4 *Restriction assays*

To check if isolated positive colonies actually contain plasmid with the sequence of interest, cloned DNA fragment was cut from plasmid using BamHI and XhoI enzymes. It is essential, that even after subcloning the sequences of interest from pGEX-4T-2 to pET42, BamHI and XhoI restriction sites still flanked the sequences. The reaction mixtures contained 5 µl of purified DNA, 0,5 µl of BamHI and XhoI enzymes (10U), 1 µl of 10x BamHI buffer, 0,1 µl of BSA, and 2,9 µl of ddH₂O was added to the final volume of 10 µl. Reactions were carried for 1 hour at 37°C. Cutting was verified by 1% agarose gel electrophoresis.

2.4.1.5 DNA sequencing

Clones, which were positive after screening of colonies and restriction assays, were given to sequencing to verify correctness inserts in plasmids. Sequencing was performed by Medigenomix or MWG – Biotech (Germany).

2.4.1.6 Cloning to LIC vectors

2.4.1.6.1 Primers

Primers with respective overhangs compatible with give vector were designed to clone constructs to LIC type vectors (Table 2.10).

Vector		Overhang sequence
pET30 Xa/LIC	sens	5' – GGTATTGAGGGTCGC – 3'
	antisens	5' – AGAGGAGAGTTAGAGCCTTA – 3'
pET41 Ek/LIC	sens	5' – GACGACGACAAGAT – 3'
	antisens	5' – GAGGAGAAGCCCGGT – 3'

Table 2.10. Primers for cloning to vectors of LIC type has to have corresponding overhangs.

2.4.1.6.2 PCR

Reaction was prepared accordingly:

5 µl	10x buffer	
0,5 µl	dNTP	100mM
50 or 100 ng	template DNA	
2,5 µl	sens_primer	10 pmol/µl
2,5 µl	antisens_primer	10 pmol/µl
water (up to 47,5 µl)		
2,5 µl	PWO polymerase from Peqlab (after initial heating)	

PCR cycler was programmed depending on the primers properties for example:

Lid 105°C

1 cycle

94°C 2 min

20 cycles

94°C 1 min

72°C 45 sec

1 cycle

72°C 7 min

4°C hold

PCR product has to be purified to remove dNTP and template plasmid before proceeding to T4 DNA polymerase digestion. PCR products were purified using QIAquick PCR Purification Kit according to the manufacturer protocol.

2.4.1.6.3 T4 DNA polymerase digestion

The following components were assemble in a sterile 1.5 ml microcentrifuge tube kept on ice:

0.2 pmol of purified PCR product in up to 14.6 µl water
 2 µl 10X T4 DNA Polymerase Buffer
 2 µl dGTP (pET30 Xa/LIC) or dATP (pET41 Ek/LIC) 25 mM
 1 µl DTT 100 mM
 0.4 µl T4 DNA Polymerase 2.5 U/µl (0.5 unit per 0.1 pmol
 PCR product) (LIC – qualified)

Nuclease free water up to 20 µl

Final concentration of insert is 0.01 pmol/µl

Concentration of PCR product was calculated from ug/ul to molar concentration using relation – number of bp in insert × 650 = pg/pmol.

Reaction was incubated at 22°C for 30 minutes and subsequently inactivated by heating up to 75°C for 20 minutes.

Ligation reaction was assemble with following components:

1 µl LIC Vector (approximately 0.01 pmol of LIC vectors)
 2 µl (0.02 pmol) T4 DNA polymerase-treated LIC insert

Reaction was incubated at 22 °C for 5 minutes and then after adding 1 µl of 25 mM EDTA again 5 minutes at 22 °C. Resulting 4 µl of ligation reaction mixture was used for transformation of E.Coli competent cells.

2.4.1.7 *Gen synthesis*

Primers for gensynthesis were designed to span desired protein sequence. They were optimized considering E.coli codon usage and primers secondary structure (calculation using Netprime). Moreover they have to overlap on at least 15 bp.

2.4.1.8 *Site directed mutagenesis*

2.4.1.8.1 *Primers*

Site directed mutagenesis requires primers which have:

1. desired mutation and overlapping sequence
2. length of at least 25 – 45 bp
3. T_m bigger or equal to 78 °C ($T_m = 81,5 + 0,41(\%GC) - 675/N - \% \text{ mismatch}$)
4. mutation positioned around the center of primer sequence and surrounded by at least 10 – 15 bp
5. optimally GC% around 40 % or higher
6. FPLC or PAGE purity
7. at the ends at least 2 G or C
8. mutation with frequent codon

To check primers for secondary structure software NetPrime from BioSoft was used (<http://www.premierbiosoft.com/netprimer/netprimer.html>)

To increase number of positive clones after mutagenesis reaction, optimized protocol was used (Wang et al. 1999; Wang et al. 2001).

2.4.1.8.2 *PCR*

For give template vector size (for example 5,5 kbp) PCR cycler was set accordingly

First step reaction

1 cycle

95°C 30 sec

2 cycles

95°C 30 sec

55°C 60 sec

68°C 11 minutes (2 minutes each 1kb)

Second step reaction**1 cycle**

95°C 60 sec

13 cycles

95°C 30 sec

55°C 60 sec

68°C 11 minutes (2 minutes each 1kb)

1 cycle

4°C 120 sec

Dpn I digestion

37°C 60 minutes

Primers working solution concentrations were 0,1 µg/µl, dNTP's 100 mM,

For first step reaction two eppendorfs one for each primer was prepared accordingly:

2,5 µl	Reaction Buffer
1,25 µl	sens_primer / antisens_primer
0,5 µl	dNTP mix
0,25 µl	(25 ng) 100 ng/ul DNA templates
SUME=4,5 µl	

sterile water	up to 24,5 µl (20 µl)
0,5 µl	<i>PfuTurbo</i> polymerase (2,5 U/µl)
SUME=25 µl	

For the second step both reactions from the first step were mixed.

2.4.1.8.3 *Dpn I digestion*

After completing second step DNA template was degraded using 1 µl of Dpn I enzyme (10 U/µl) to 50 µl. Digestion at 37°C was incubated for 60 minutes. For subsequent E.Coli transformation 1 µl of reaction mix was used.

2.4.1.8.4 Transformation

Transform 5ul of the reaction product into Top10/XL-Blue or Nova-Blue cells as per standard protocol, and plate on the LB-agar plate with the desired antibiotic.

2.4.1.8.5 Selection of the mutant

Make the mini-prep plasmid isolation of 3-10 colonies and sequence the clone with the desired primer to select the mutant.

2.4.1.9 In vitro protein synthesis

Solutions of separate amino acids in water were prepared according to Table 2.11. It was required to decrease pH of some of them (Asn, Asp, Glu, Gln). To ensure good dissolving shaking and high temperature was applied. Cys and His were always prepared freshly.

Amino acid	Mw (Da)	Molar concentration (mM)	Volume (ml)	Amount (mg)
Ala	89,1	10,50	30,00	28,1
Arg	174,2	10,50	30,00	54,9
Asn-H ₂ O	150,14	10,50	30,00	47,3
Asp	133,11	10,50	30,00	41,9
Cys	121,16	42,00	1,00	5,1
Glu	147,13	10,50	30,00	46,3
Gln	146,15	10,50	30,00	46,0
Gly	75,05	10,50	30,00	23,6
His	156,15	40,00	1,00	6,25
Ile	131,18	10,50	30,00	41,3
Leu	131,18	10,50	30,00	41,3
Lys-HCl	182,65	10,50	30,00	57,5
Met	149,21	10,50	30,00	47,0
Phe	165,19	10,50	30,00	52,0
Pro	115,13	10,50	30,00	36,3
Ser	105,09	10,50	30,00	33,1
Thr	119,12	10,50	30,00	37,5
Trp	204,23	10,50	30,00	64,3
Tyr	181,19	10,50	30,00	57,1
Val	117,15	10,50	30,00	36,9

Table 2.11. Concentrations of amino acids used for RTS – 500.

Protein was synthesized *in vitro* using RTS – 500 (ROCHE). Protocol was adapted from manufacture manual. Selective labeling of protein was possible using this kit and amino acids mix prepared according to labeling strategy.

810 µl	Amino acids mix
150 µl	¹⁵ N – His
150 µl	Cys
1890 µl	Reconstitution Buffer (provided by manufacturer)

Reaction was set up accordingly to protocol 3.3.1.2 from the page 16 of RTS – 500 manual. Parameters were used rpm – 120; temperature – 30°C; time – 24 hours.

Reaction mixture was centrifuged and buffer was changed to buffer A on gel filtration column (Sephadex G25) [Amersham Pharmacia Biotech]. GFPuv was purified by immobilized metal affinity chromatography (IMAC) (Ford et al. 1991) on a Ni-NTA (nickel-nitrilotriacetic acid) column [Qiagen] according to manufacture procedure (as a elution buffer B was used). Protein was concentrated and dialyzed against buffer C.

Buffer	Composition
Buffer A	50 mM NaH ₂ PO ₄ , 300 mM NaCl, 10 mM imidazole, pH 8
Buffer B	50 mM NaH ₂ PO ₄ , 300 mM NaCl, 250 mM imidazole, pH 8
Buffer C	8 mM KH ₂ PO ₄ , 16 mM Na ₂ HPO ₄ , 112 mM NaCl, pH 7

Table 2.12. Buffers used for *in vitro* synthesized protein.

2.4.1.10 *Escherichia Coli* transformation

All *E. coli* competent cells strains were from Invitrogen (Netherlands) (unless otherwise mentioned). Transformation was performed according to protocol attached to competent cells from Invitrogen.

2.4.1.10.1 Chemical transformation

2.4.1.10.1.1 PROTOCOL FOR MAKING CHEMICALLY COMPETENT CELLS

Solutions required:

1. CaCl₂·2H₂O (1 M); Filter sterilized (0.22 U)
2. MgCl₂-CaCl₂ solution (80 mM MgCl₂; 20 mM CaCl₂); Filter sterilized (0.22 U)

Pick a single bacterial colony from a plate that has been incubated for 16-20 hours at 37⁰. Transfer the colony into 100ml of LB broth in one liter flask. Incubate the culture for 3 hours at 37⁰C with vigorous agitation, monitoring the growth of the culture. Grow the cells till the OD600 comes to around 0.40.

Transfer the bacterial cells to sterile, disposable, ice-cold 50-ml polypropylene tubes. Cool the cultures to 0⁰C by storing the tubes on ice for 10 minutes.

2.4.1.10.2 Electro transformation

2.4.1.10.2.1 PROTOCOL FOR ELECTROCOMPETENT CELLS:

1. Bacteria were streaked on an LB agar plate, and incubated at 37⁰C overnight.
2. 50 ml of LB medium in a 250 ml flask were inoculated with a single colony from the LB plate and incubated at 37⁰C with shaking (200 rpm) overnight.
3. 1 l of LB medium in a 3l flask was inoculated with the 50ml overnight culture. The culture was grown in shaking (200 rpm) incubator at 37⁰C until the OD600 was between 0.5 – 0.6 (approximately 2 hours).
4. The culture was transferred to the two chilled, sterile 500ml centrifuge bottles and incubated on ice for 30 min. Thereafter centrifugation followed at 2000 G for 15 minutes at 0 – 4⁰C.
5. Supernatant was decanted, and bottles put back on ice. The cell pellet in each bottle was resuspended in approximately 500 ml of cold (0 – 4⁰C) sterile water, and subsequently centrifuged like before.
6. The cells in each bottle were washed again with 250 ml of cold sterile water, and centrifuged.
7. The cell pellet in each bottle was then resuspended in 20 ml cold sterile 10 % glycerol and transferred to a chilled, sterile, 50 ml centrifuge tube. Centrifugation followed at 4000 G for 15 minutes at 0 – 4⁰C.
8. The 10 % glycerol was decanted and pellet resuspended for the second time in 1ml cold sterile 10 % glycerol.
9. Using a pre-chilled pipette the cell suspension was aliquoted (40 µl) to pre-chilled 1.5ml tubes and frozen immediately in liquid nitrogen. The aliquots were kept at – 80⁰C ready for use.

2.4.1.10.2.2 ELEKTROPORATION

1µl of plasmid DNA solution in water was mixed together with the 40µl aliquot of electrocompetent bacteria and put between the electrodes of a 0.1 cm electroporation cuvette (Biorad, FRG). The cuvette was then put into the electroporator (Stratagene, FRG), and a pulse of 1660 V was applied. The value of the time constant was observed (usually 3-5 ms). The mixture was then washed out from between the electrodes with 1 ml of sterile pre-warmed (37°C) LB medium (without antibiotics), transferred to a sterile 1.5 ml tube and shaken (800 rpm) at 37°C. After 1 hour the cells were streaked on a LB agar plate with an appropriate antibiotic.

2.4.2 *Bacterial culturing*

2.4.2.1 *Bacterial culture in LB medium:*

1. 15ml LB were inoculated with a fresh single bacterial colony and incubated overnight at 37 °C with vigorous shaking (280 rpm) in 50 ml falcon tube.
2. Pre-warmed 1 l LB medium in 3 l flask was inoculated with 10 ml of the overnight culture, supplemented with appropriate antibiotic, and incubated at 37°C with Shaking (180 rpm) until the OD600 reached the 0.7-0.8 value.
3. Induction by addition of 1 mM IPTG (final concentration) was usually used. The cells were then grown until the expected OD was reached.
4. Cells were pelleted by centrifugation at 4200 rpm in Beckman centrifuge for 40 minutes at 4°C.
5. Supernatant was discarded and pellet was stored at -70°C.

2.4.2.2 *Bacterial culture in MM:*

1. 2 ml LB were inoculated with a single colony, and shaken (150 rpm) overday in 15 ml falcon tube at 37°C.
2. 20 ml MM were inoculated with 50 µl the overday culture, and shaken (280 rpm) overnight in 100 ml flask at 37°C.
3. 1 l MM was inoculated with 20ml the overnight culture (1:50), and shaken (150 rpm) in 3 l flask, until the expected optical density was reached.

4. Cells were pelleted by centrifugation at 4200 rpm in Beckman centrifuge for 40 minutes at 4°C
5. Supernatant was discarded and pellet was stored at -700C until used.

2.4.2.3 *Medium for selectively enrichment amino acids*

Respective chemicals were mixed:

Ala, Glu, Gln, Arg, Gly 400 mg

Asp, Met 255 mg

Cytosine, Guanosine, Uracil 125 mg

Asn, Ile, Val, Leu, His, Lys, Pro, Thr, Tyr 100 mg

Phe, Thymidine, Thymine 50 mg

CaCl₂ 10 mg

NaAcetate 2 g

KH₂PO₄ 10 g

Citric acid 1 g

Trace elements solution 1 ml

Ferrous citrate (preparation see protocol for MM) 36 mg

Zn-EDTA 1 ml

Ser 1.6 g

NH₄Cl 1 g

The pH was adjusted to 7.4 with 1 M NaOH,

The mixture was autoclaved, after autoclaving the following sterile filtered solutions were added:

Cys, Trp, nicotinic acid 50 mg

Biotin 0.1 mg

Mg₂SO₄ 2 ml

Thiamin 560 µl

Glucose 20 ml

Antibiotics (2x of the usual amount for the LB-medium)

Concentrations of stock solutions' were like for MM.

2.4.3 *Protein production techniques*

2.4.3.1 *Protein expression in Escherichia Coli strains*

2.4.3.1.1 *Cell differentiation and cell cycle proteins*

Id-2, MyoD and pRb constructs were expressed in Escherichia Coli respectively BL21 STARTM (DE3) and BL21 STARTM (DE3) pLysS (Invitrogen) strains. Cells freshly transformed with given MyoD construct were grown in LB with kanamycine at 37^o Celsius up to OD600 reached 0.7. After IPTG was added to the end concentration of 1mM culture was shaken for additional 4 hours in 37 °C. Cells freshly transformed with His - ABC construct were grown in LB with kanamycine and 100uM MgCl₂ at 37 °C up to OD600 0.4 then temperature was decreased to 18 °C. When the culture reached OD600 0.8 IPTG was added to final concentration 1mM. Induction time was approximately 10 hours.

BL21 STARTM (DE3) pLysS cells freshly transformed with ABC-C or ABC-CL construct were grown in LB supplemented with 100 µg/ml ampicillin, 34 µg/ml chloramphenicol 1% Glucose 10 mM K₃PO₄ pH 7.2 and 1 mM MgSO₄ at 37 °C up to OD600 0.3 then temperature was decreased to 21 °C. When the culture reached OD600 0.8 IPTG was added to final concentration 1mM and temperature was decreased to 14 °C. Induction time was approximately for 12 hours.

To obtain uniformly or selectively labeled samples M9 minimal medium supplemented with respectively ¹⁵N-NH₄Cl, ¹⁵N-Lys, ¹⁵N-Leu or others amino acids were used.

2.4.3.1.2 *Expression of labeled samples of GFPuv*

¹⁵N, ¹⁵N/99%-²H, ¹⁵N/¹³C, and ¹⁵N/¹³C/70%-²H labeled samples of GFPuv were produced as described by Georgescu et al., (2000 and 2002). The samples were dissolved in phosphate buffered saline (PBS) at pH 7.0. Expression was controlled by the T7 promotor and provided by the gene for β-lactamase. After induction with 0.5 mM isopropylthiogalactoside (IPTG) His6-tagged GFP was overexpressed in BL21(DE3) cells in 1 liter LB-medium for 6 h at 37°C. The cell pellet was resuspended in 50 ml lysis buffer (300 mM NaCl, 50 mM Na₂HPO₄, pH 8.0) with 1 mg DNase, 1 mg RNase, 1 mg MgCl₂

and 0.1 M phenylmethylsulfonyl fluoride (PMSF) and then sonicated with Micro-Tip and centrifuged for 45 minutes at 80.000 rpm in a Beckmann centrifuge.

2.4.3.1.3 Expression of fluoro tryptophane GFPuv

The incorporation of fluorinated analogues 4-fluoro-Trp (4FW), 5-fluoro-Trp (5FW) and 6-fluoro-Trp (6FW) into Clontech GFPuv (F99S, M153T, V163A), Clontech EGFP (F64L, S65T) and Clontech ECFP (F64L, S65T, Y66W, N146I, M153T, V163A) was achieved in *E. coli* Trp-auxotrophic strain ATCC49980 (Minks et al., 1999) using the selective pressure incorporation (SPI) method developed in our lab (Minks et al., 2000). The protein expression host *E. coli* ATCC49980 was routinely co-transformed with two plasmids: ampicillin resistant pQE-30-PP4 harboring the GFP and CFP gene sequence under control of the T5 promoter, and kanamycin resistant pREP4 containing a repressor gene lacIq. The native and fluorinated proteins included a polyhistidine tag and were purified using Ni-chelate columns.

2.4.3.2 Expression in Spodoptera frugiperda strain SF9

CDK6 and CDK4 – cyclin D constructs were cloned between BamHI and EcoRI sites into pAcG2T vectors. Infective viruses were produced and right titer was checked. Protein expression was performed from 4000 ml (cell density 1mln/1000ml) at 27⁰C. After 58 hours cell were spun and frozen.

2.4.4 Protein purification techniques

2.4.4.1 Cell differentiation and cell cycle proteins

2.4.4.1.1 Protein purification

All pRb constructs were passed in the first step of purification through NiNTA (Qiagen). Lysis buffers consist of 50 mM NaH₂PO₄ pH 8, 300 mM NaCl, 10 mM BME 10mM of the imidazole. Additionally the lysis buffer was supplemented with 1 tablet to 50 ml of protease inhibitor set “Complete” without EDTA (Roche). The sample was loaded onto the column and washed with lysis buffer supplemented with 20 mM imidazole. Elution was performed with lysis buffer supplemented with 250 mM imidazole. The ABC–C and

ABC-CL constructs were subjected to the MonoQ (Amersham-Pharmacia) anionic exchange chromatography. Buffer A consists of 50 mM NaH₂PO₄ pH 7.8 10 mM BME buffer B had additionally 1 M NaCl.

Id-2 and, MyoD constructs were purified over NiNTA (Qiagen). Buffers consists of 6 M GuaHCl, 50 mM Tris (lysis pH 8; wash pH 6 and pH 5.2; elution pH 4.5). Refolding was made by dialysiation of low concentrated protein sample to 50 mM NaH₂PO₄, 300 mM NaCl, pH 7.8, 10 mM BME. Sf9 were lysed in a lysis buffer (50 mM Na₂HPO₄ pH8, 250 mM NaCl 5mM BME), sonicated and span.

CDK6 was purified over glutiation-sepharose (Amersham) according manual. To remove GST from CDK6 the thrombin cleavage was performed (0.25 U/ml) on the glutiation-sepharose column. Protein was further purified to homogeneity on S75 Sephadex.

2.4.4.1.2 DAPzyme digestion

Prior to a pull-down assay, his-tag was removed from the pRb samples by DAPzyme (Qiagen) digestion (reaction was performed according to manufacturer manual) at room temperature overnight. Proteins were subsequently purified over NiNTA to remove His-tags or undigested proteins. His-tag removal was controlled using SDS-PAGE and Western-blot.

Identity of all constructs was checked on the DNA level by DNA sequencing. All proteins were checked by Western-blot, mass-spectroscopy and/or N-terminal Edman-sequencing. Purity of all proteins was estimated from SDS-PAGE and mass spectrometry to be 95-98%.

2.4.4.2 GFPuv and mutant proteins

All labeled samples of GFPuv were purified by Ni²⁺ affinity chromatography. Ni-NTA resin (Qiagen) was incubated with the supernatant for 1 hour at 4°C, poured into a column and subsequently washed with 100 ml of a lysis buffer. The protein was eluted from the Ni-NTA resin with 20 ml elution buffer (300 mM NaCl, 50 mM Na₂HPO₄, 500 mM imidazole, pH 8.0). Then the protein was dialyzed against PBS (115 mM NaCl, 8 mM KH₂PO₄, 16 mM Na₂HPO₄, pH 7.0) and concentrated to 15-20 mg/ml using Centricon 10 (Milipore). After additional gel filtration with a Superdex 75 column (Pharmacia), the

sample was concentrated again to a concentration of 20-25 mg/ml. The purity of the sample was checked by mass spectroscopy.

2.4.5 Handling and storing of the proteins

All the operations involving the proteins at either level of purification were performed on ice or at 4-50C, unless denaturing conditions were used. A stress was laid on speed of processes and decrease of time consumption and number of purification steps. 2-mercaptoethanol was added to all buffers to prevent disulfides bond formation. PMSF was added shortly before every run, as it is quickly inactivated in aqueous solutions. Complete inhibitors cocktail EDTA-free (Roche, Germany) was also added. It was present to protect protein of interest against proteolysis. The function of NaN₃ was to prevent bacterial and fungal growth in solutions stored for prolonged time. Purified protein was stored at +4oC or at -80oC to further analysis.

2.4.6 Analytical methods

2.4.6.1 Protein detection

2.4.6.1.1 SDS-PAGE

Results of overexpression tests, levels of protein production as well as quality and purity of fractions obtained during successive steps of purification were analyzed by discontinuous sodium dodecyl sulphate-polyacrylamide gel electrophoresis. Mini-gels (70x80x0.75 mm) were electrophoresed in Bio-Rad MINIPROTEAN II apparatus. The Tris-Tricine SDS-PAGE was adapted from Schagger & von Jagow (1987).

Separating gel composition I (17 %)

ADDW	1.7 ml
3 M Tris, pH 8.8	2.5 ml
Acrylamide	2.5 ml
Glycerol	0.8 ml
APS(10%)	25 µl
TEMED	2.5 µl

Add 2.25 ml per gel.

Stacking gel composition II

ADDW	1.7ml
3 M Tris, pH 8.8	1.3 ml
Acrylamide	0.75 ml
APS (10%)	2.5 μ l
TEMED	1.25 μ l

Add 1.5 ml per gel

Different buffers were used for anode and cathode:

Anode buffer (+)

0.2M Tris-HCl, pH 8.9 24.22 g/l

Cathode buffer (-)

0.1M Tris-HCl, pH 8.25 12.2 g/l

0.1M Tricine 17.9 g/l

0.1 % SDS

Samples of cells to be analyzed for protein expression were prepared by resuspending the bacterial pellet, from 1 ml volume of culture OD_{600nm}= 1, in 100 μ l of 5x sample buffer and boiled for 5 min.

For visualization of separated proteins, gels were stained with Coomassie Brilliant Blue R-250 by incubation of the gels in staining solution for 0.5 h, followed by 2-3 h destaining in destainer solution. Overnight incubation led to contrast increase and swelling of the gels, what made interpretation of results easier. The gels were scanned on EPSON 2100.

2.4.6.1.2 *Western immunoblotting*

This method of protein detection was applied to confirm immunological identity of the expressed proteins and to determine presence of any degradation products or isoforms of the purified proteins.

The proteins separated by SDS-PAGE were blotted in semi-dry blotter (home made). One 1 mm filter paper soaked in blotting buffer was placed on the bottom anode plate of the apparatus. A nitrocellulose membrane (Amersham Bioscience), SDS-PAGE and one filter paper soaked in blotting buffer were placed on the top of it. The blotting was carried out at 200 mA for 1 hour. The NC membrane was then immersed in 20 ml of 5% milk powder

solution in PBS-T and gently agitated for 1/2 h, then respective primary antibody was added and NC membrane was agitated for additional 1 h. The membrane was washed 3x 15 min with 20 ml of washing buffer and incubated for 1 h with secondary antibody – AP-conjugated goat anti-mouse immunoglobulins (1:2000 dilution) (Santa Cruz) 5% milk powder solution in washing buffer. The NC membrane was washed 3x5 min with 20 ml of water and for visualization substrate (SIGMA FAST BCIP/NBT) for AP was added (Sigma) for 10 minutes.

Transfer buffer

25 mM TrisHCl pH 8.3

192 mM Glycine

To make the final working solution mix 80ml of buffer with 20ml of Methanol

Alkaline Phosphatase Buffer

100 mM TrisHCl pH 9.5

100 mM NaCl

5 mM MgCl₂

Wash Buffer

10 mM TrisHCl pH 8.0

150 mM NaCl

0.05 % Tween20

Substrate for alkaline phosphatase (BCIP)

Dissolve 1 tablet in 10ml of water (as directed by manufacturer, SIGMA)

2.4.6.1.3 Determination of protein concentration

The concentration of proteins in solutions was estimated using the Bradford reagent (Biorad; Bradford, 1976). 5µl of the protein solution (or 1µl, if the protein solution is very concentrated) to be measured were added to 495 µl of Biorad-reagent working solution (working solution was prepared by 1:10 dilution of Biorad-reagent stock solution in PBS buffer or water, stored in the fridge). After thoroughly mixing the sample, the OD595 was measured. As a reference similar mixture was prepared with 5µl buffer instead of protein

solution. OD was subsequently converted into the protein concentration on the basis of a BSA calibration curve.

2.4.6.2 Pull down assays

The “bait” proteins with the His-tag or GST-tag were bound to the respective columns. Prey with enzymatic cleaved tag was put over column saturated with the bait. Complex was eluted with 250 mM imidazole or 10 mM reduced glutathione. Fractions obtained were tested on SDS-PAGE.

2.4.6.3 Analytical gel filtration

A mixture of two proteins was subjected to separation on the Sephadex S75 (Amersham-Pharmacia) column. Fractions were checked on SDS-PAGE and after concentrating with mass spectroscopy.

2.4.6.4 Mass spectroscopy

Mass spectroscopy was carried out on ESI-MS API 165 Perkin-Elmer Sciex (Langen) coupled with HPLC (column: Macherey-Nagel EC 125/2 Nucleosil 300-5 C4 MPN; pump system: Microgradient System 140B/C Perkin Elmer (solvent A: water, 0.05 % TFA, B: MeCN, 0.05% TFA; gradient 10-95% B); photodiode array Agilent HP1100PDA; software: Masschrom, Biomultiview).

2.4.6.5 Functional assays of cell cycle and differentiation proteins

Folding of proteins was checked by 1D-proton NMR spectra. MyoD constructs were additionally checked by functional DNA-binding assays and Id-2 was checked for formation of a complex with MyoD. The pRb was fully functional as proved by the interaction with the E7 full-length protein, and peptides (E7-1, E7-0), the SV40 large T antigen peptide, and the E2F construct spanning residues 243-437. CDK6 was active as proved by binding to p19.

2.4.6.6 *Optical Spectroscopy*

2.4.6.6.1 *GFPuv*

UV-absorption spectra of 10 μM GFPuv and GFPuv His148Gly in PBS buffer were recorded on a Perkin-Elmer UV/VIS spectrophotometer at room temperature. The spectra were scaled using a known molar extinction coefficient of GFP ($\epsilon\text{M} = 21000 \text{ M}^{-1} \text{ cm}^{-1}$ at $\lambda = 277 \text{ nm}$). Optical excitation and emission spectra were measured on a Perkin-Elmer LS50B spectrometer at RT. The protein concentration was adjusted to 50 nM. In order to acquire the excitation spectra, the emission at 510 nm was measured. To record the emission spectra, the samples were excited at 396 nm. In both cases a slit of 2.5 nm was used.

2.4.6.6.2 *Fluoro tryptophane GFPuv*

UV-absorption spectra of the proteins in the phosphate buffered saline (PBS: 115 mM NaCl, 8 mM KH_2PO_4 , 16 mM Na_2HPO_4 , pH 7.3) were recorded on a Perkin-Elmer Lambda 17 UV/VIS spectrophotometer. Molar extinction coefficients (ϵM) for native and fluorinated proteins were determined at room temperature (20 °C). Native and 6FW-EGFP have essentially the same absorbance ($\lambda_{\text{max}} = 277 \text{ nm}$: $\epsilon\text{M} = 21000 \text{ M}^{-1} \text{ cm}^{-1}$; $\lambda_{\text{max}} = 488 \text{ nm}$: $\epsilon\text{M} = 38000 \text{ M}^{-1} \text{ cm}^{-1}$). Molar extinction coefficients (ϵM) for native and substituted ECFP variants were different: native ECFP: $\lambda_{\text{max}} = 278 \text{ nm}$: $\epsilon\text{M} = 25000 \text{ M}^{-1} \text{ cm}^{-1}$; $\lambda_{\text{max}} = 434 \text{ nm}$: $\epsilon\text{M} = 25500 \text{ M}^{-1} \text{ cm}^{-1}$; 6FW-ECFP: $\lambda_{\text{max}} = 280 \text{ nm}$: $\epsilon\text{M} = 23300 \text{ M}^{-1} \text{ cm}^{-1}$; $\lambda_{\text{max}} = 430 \text{ nm}$: $\epsilon\text{M} = 22900 \text{ M}^{-1} \text{ cm}^{-1}$. Fluorescence spectra were recorded on a Perkin-Elmer spectrometer (LS50B) equipped with digital software. Protein samples (0.25 μM ; slit 2.5 nm) were excited at 488 nm (EGFP) or at 450 nm (ECFP) and the emission spectra were recorded in the 490 – 540 nm (EGFP) or 460 – 540 nm (ECFP) range. Emission and excitation spectra of 6FW ECFP were measured at temperatures of 288K, 293K, 298K, 303K, 308K, 313K, 318K, and 323K.

2.4.6.7 *NMR*

All spectra were recorded on the 500, 600, 750, 900 MHz Bruker spectrometer equipped with a 5 mm ^1H - ^{13}C - ^{15}N triple resonance TXI probeheads, including triple-axis gradients (DRX600, DMX750) or z-axis gradients (DRX500). The GFPuv spectra were recorded at 310K with a sweepwidth of 17.5 ppm for ^1H and 42 ppm for ^{15}N . The GFPuv assignment of the backbone resonances was accomplished with triple resonance spectra

including HNCA (Cavanagh et al., 1996), TROSY-HNCA (Salzmann et al., 1998), HN(CO)CA, HNCO, CBCA(CO)NH (Cavanagh et al., 1996). A HBHA(CBCACO)NH (Cavanagh et al., 1996), spectrum was also recorded on a Bruker Cryoprobe™ DRX 600 spectrometer. In addition, HSQC (Mori et al. 1995) spectra of selectively ^{15}N labeled Gly, Phe, Tyr, Met, Ala, Leu, Lys, Ile, and Val and inverse ^{15}N Asn, Thr, His, and Asp labeled samples of GFPuv were recorded. 2D NOESY ($\tau_m=160\text{ms}$) spectra and 3D NOESY-HSQC ($\tau_m=160\text{ms}$) spectra were recorded. Where appropriate, the spectra were recorded on 70% or 99% deuterated samples of GFPuv. All spectra were processed using XWINNMR (Bruker). Details of the NMR data acquisition and resonance assignment for GFPuv are described in Georgescu et al. (2003).

The GFP pH titration and a proton to deuter exchange (Dempsey 2001) experiment were conducted based on ^1H - ^{15}N HSQC spectra (Mori et al. 1995). For the titration the pH of a ^{15}N labeled sample of GFPuv in PBS buffer was reduced from pH 7.7 to pH 6.1 in series of 0.1 by adding the appropriate amount of 1M H_3PO_4 . For each step a HSQC spectrum was recorded. For the HD exchange experiment the ^{15}N labeled sample of GFPuv was lyophilized and dissolved in 99.99% D_2O . Then a series of 16 HSQC spectra were recorded. The measurement time for one spectrum was 1 hour. In the following month one spectrum per week was recorded. Finally spectra after 2, 3 and 5 months were recorded, between measurements the protein was stored at $+4^\circ\text{C}$.

^{19}F NMR measurements were carried out on a Bruker DRX500 spectrometer equipped with a dual ^1H - ^{19}F probehead. The fluorine Larmor frequency was 470 MHz. All measurements were performed at a temperature of 303 K with exception of six spectra taken at temperatures between 293 K and 318 K. Due to the negative NOE enhancement factor of fluorine nuclei in proteins no proton decoupling was used.

2.4.6.8 NMR diffusion measurements

The translational diffusion coefficient (Dingley et al. 1995) of GFPuv at 310 K was measured by a modified Watergate-1H-1D-sequence as described by others (Sklenar et al. 1993). The gradients were calibrated by 1D NMR imaging of a 3.5 mm Teflon plug in a standard 5 mm NMR tube filled with CuSO_4 doped water. The signal attenuation I/I_0 due to translational diffusion is given by:

$$I = I_0 e^{-bD}$$

with D being the translational diffusion coefficient. The corresponding b-value for sineshaped gradients was calculated using the relationship described at Seifert et al. (2002).

$$b = \gamma^2 G \frac{\delta^2}{\pi^2} (4\Delta - \delta)$$

where G represents the maximum strength of the gradient. The apparent mass of the protein can be calculated from the diffusion coefficient according to (Cantor 1980):

$$M = \left(\frac{kT}{6\pi\nu FD} \right)^3 \frac{4\pi N_A}{3(V_2 + \kappa V_1)}$$

For the apparent mass calculation the following parameters (T = 310K) were used: the protein form factor F = 1.03, hydration of the protein $\kappa = 0.34$, viscosity of water $\nu = 6.91 \times 10^{-4}$ Ns/m², specific volume of the protein $V_2 = 0.731 \times 10^{-3}$ m³/kg, and specific volume of water $V_1 = 1.00669 \times 10^{-3}$ m³/kg. k and N_A represent the Boltzmann constant and Avagadro's number, respectively.

2.4.6.9 NMR relaxation measurements

For investigation of the backbone flexibility, the relaxation parameters T₁ (Skelton et al. 1993) and T₂ (Zweckstetter et al. 1998), the heteronuclear ¹H-¹⁵N Overhauser effect (Neuhaus 2000), and the ¹H-¹⁵N dipolar-¹⁵N CSA cross-correlation rate (Tjandra et al. 1996) were measured as described in the corresponding references. The measurements were performed on the ¹⁵N-labeled sample, as well as on the ¹⁵N- and ²H (99%)-labeled samples of GFPuv, at proton frequencies of 500 and 600 MHz. In the heteronuclear ¹H-¹⁵N Overhauser experiment, saturation of the amide protons was achieved by application of 120° pulses prior to the experiment (Markley et al. 1971). NOE values were given by the ratios of peak heights in the experiments with and without proton saturation. Water suppression during the T₂, NOE, and CSA relaxation experiments was achieved by using the Watergate 3-9-19 or improved Watergate 5 pulses (Sklenar et al. 1993). The relaxation delays used for determination of the T₁ relaxation times were 12.4, 260.4, 508.4, 756.4, and 1004.4 ms and for T₂ relaxation times 20.8, 41.6, 124.8, and 166.4 ms. In the T₁ experiment, water saturation was avoided by using low-power water-flip-back pulses. The recycle delay (TR) was set to 1.5 s. The T₁ measurement was conducted as described by other (Sklenar et al.

1987). The experimental peak intensities of the T_1 and T_2 experiment were fitted to the equation $A \exp(-t/T_x)$ ($x = 1$ or 2) with a Levenberg-Marquardt algorithm using the Sparky relaxation extension. The statistical errors of curve fitting were determined by Monte Carlo simulation. In the CSA experiment, the cross-correlation rate η was calculated using the relationship $I_A/I_B = \tanh(2\Delta\eta)$ with I_A and I_B being the peak intensities resulting from the two experiments described by Tjandra et al. (1996) and the delay Δ being set to 22 ms. The measurement error was derived from the signal-to-noise ratio of the corresponding peaks.

2.4.6.10 Other methods

All biochemical methods that are not mentioned here were performed exactly according to Sambrook & Russell (2001).

2.4.7 Bioinformatics und calculations

List of programs and web-pages used to analyse sequences and, structures in the supplementary materials.

3 RESULTS AND DISCUSSION

3.1 Cell differentiation and cell cycle proteins

3.1.1 The HLH domain of MyoD or Id-2 does not interact with the pocket domain of pRb

3.1.1.1 Results

All the proteins used for experiments were purified to 95-98% as examined by SDS – PAGE and mass spectroscopy. Efficiency of DAPzym digestion was tested with both affinity chromatography and Western-blot (Figure 3.1).

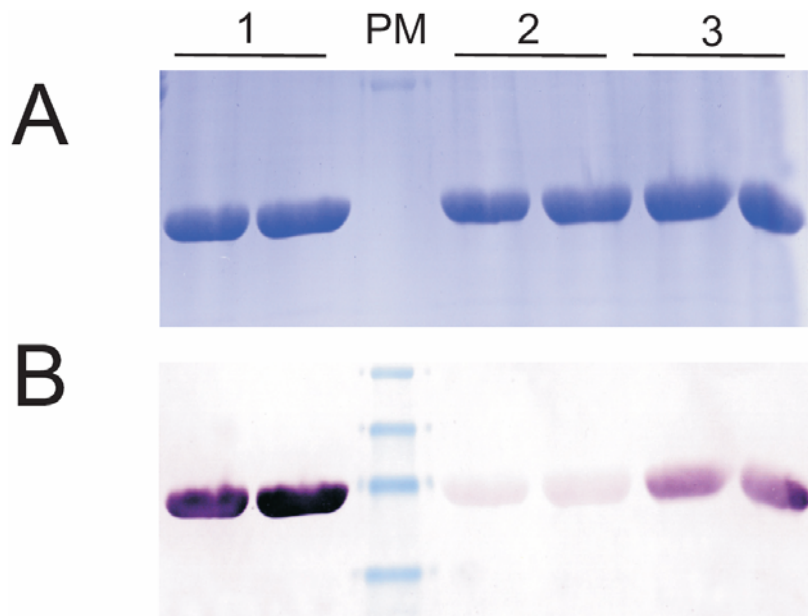


Figure 3.1. After DAPzyme reactions efficiency of digestion was tested with Western-blot. (A) Proteins stained with Coomassie blue. (B) Western-blot with anti His-tag antibodies: 1 – undigested protein (pRb (ABC-C)); 2 – protein after overnight digestion; 3 – protein after 4 hours digestion with DAPzyme (all reaction as described in the Materials and Methods carried out at room temperature).

We observed no interaction between the pocket region of pRb (constructs ABC-C, ABC-CL, AB) and the bHLH part of MyoD using both pull down assays (Figure 3.2) and

gel filtration (Figure 3.3). Mass spectrum of concentrated samples from fractions corresponding to the pRb peak also showed no MyoD.

A pull down assay in which MyoD (construct His–MCNC) was a “bait” and pRb (construct ABC–C) a “prey” showed no binding (Figure 3.2). A reverse experiment gave the same results (Figure 3.4). The same was observed in the pull down assay where Id–2 (His–R2B) was a “bait” and pRb (construct ABC–CL) a “prey” (Figure 3.5).

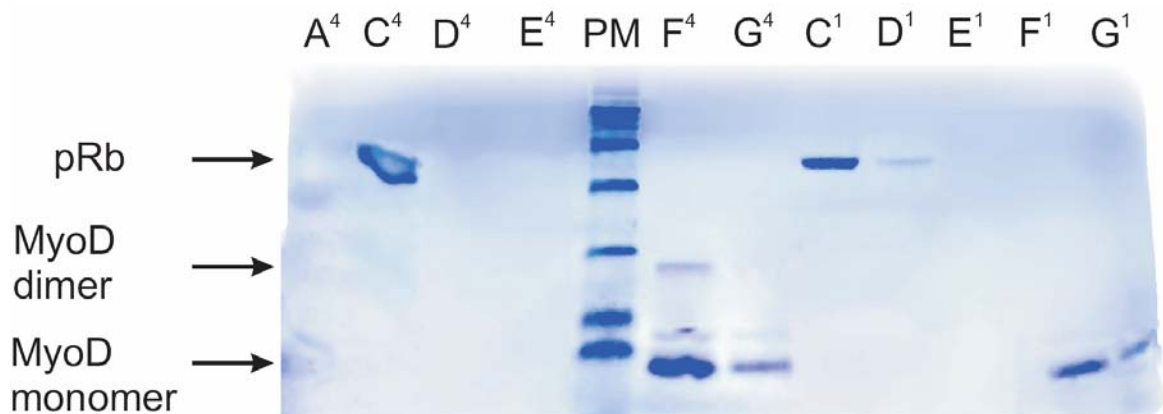


Figure 3.2. The SDS–PAGE of pull–down–assays. The His–tagged MyoD (MCNC) (“bait”) was bound to the NiNTA column and then pRb (ABC–C) (“prey”) was passed through the column. A⁴ – MyoD flow–through; C⁴ and D⁴ – pRb flow–through; E⁴, F⁴ and, G⁴ – various fractions of 250 mM imidazole. C¹, D¹ and E¹ – pRb flow through (amount of – MyoD decreased 4 times); F¹ and, G¹ – 250 mM imidazole elutions. Between lines E⁴ and F⁴ is the protein marker.

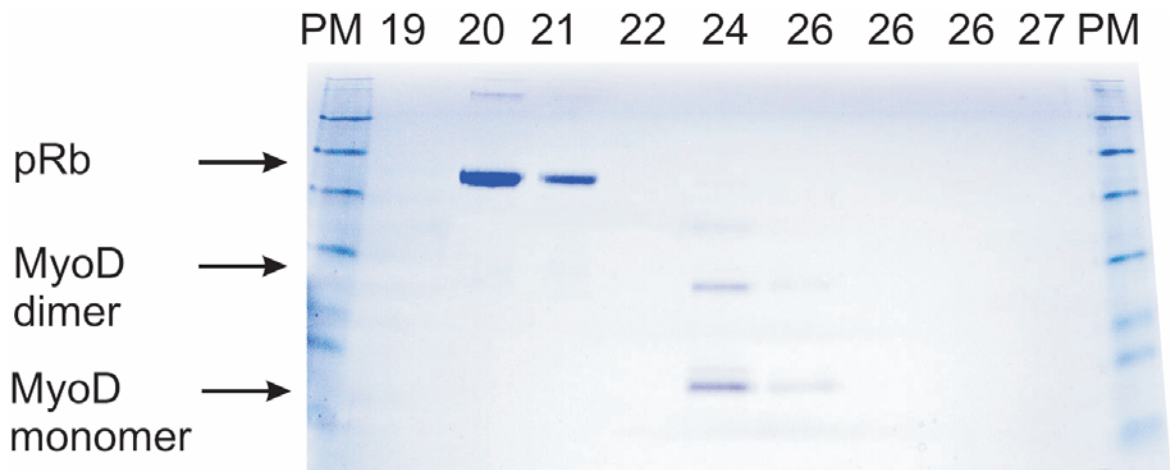


Figure 3.3. SDS-PAGE of fractions from the S-75 Sephadex gel filtration analysis of MyoD (MCNC) and pRb (ABC-C). MyoD does not coelute with pRb. There is no interaction between MyoD and pRb.

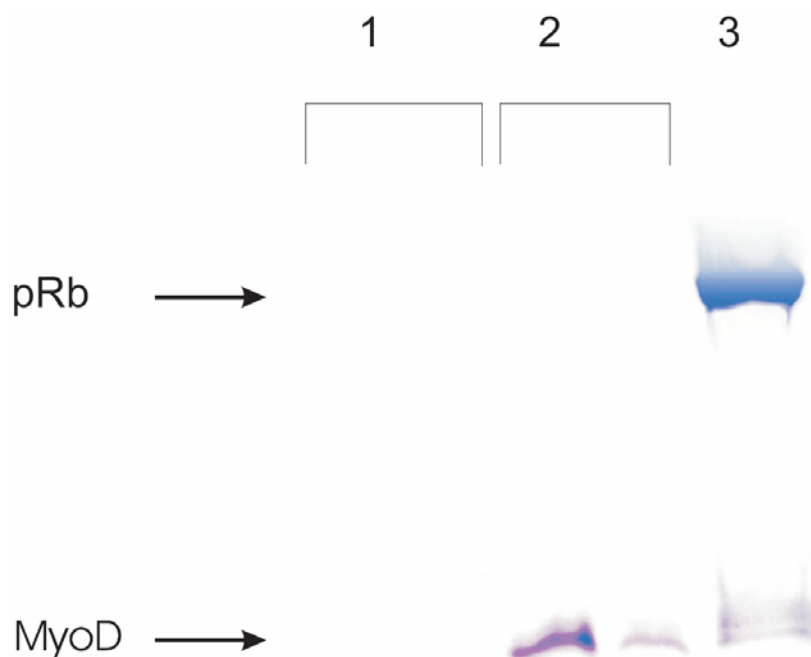


Figure 3.4. The SDS-PAGE of pull-down-assays. The His-tagged pRb (ABC-C) (“bait”) was bound to the NiNTA column and MyoD (MCNC) (“prey”) was passed through the column. 1 – pRb flow-through; 2 – MyoD flow-through; 3 – fraction of 250 mM imidazole elution.

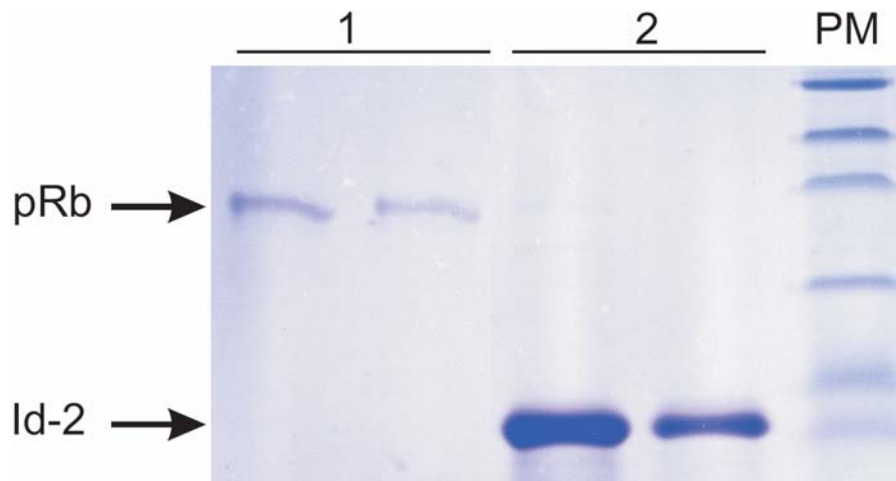


Figure 3.5. The SDS-PAGE of pull-down-assays. The HIS-tagged Id-2 (R2B) (“bait”) was bound to the NiNTA column and then pRb (ABC-CL) (“prey”) was passed through the column. 1 – pRb (ABC-CL) flow-through; 2 - 250 mM imidazole elution fraction. PM is the protein marker.

It has been reported that MyoD, Id-2 and, viral HPV16 E7 bind to the same region of pRb (Gu et al. 1993; Lasorella et al. 1996). We tested these interactions using “selectively ^{15}N labeled” pRb (^{15}N -Lys, ^{15}N -Leu). Addition of MyoD or Id-2 did not cause any changes in the ^{15}N -HSQC spectrum, proving no interaction (Figure 3.6 and Figure 3.7). In the positive control we were able to see clearly the interaction of selectively labeled samples of pRb with the E7-0, E7-1 peptides.

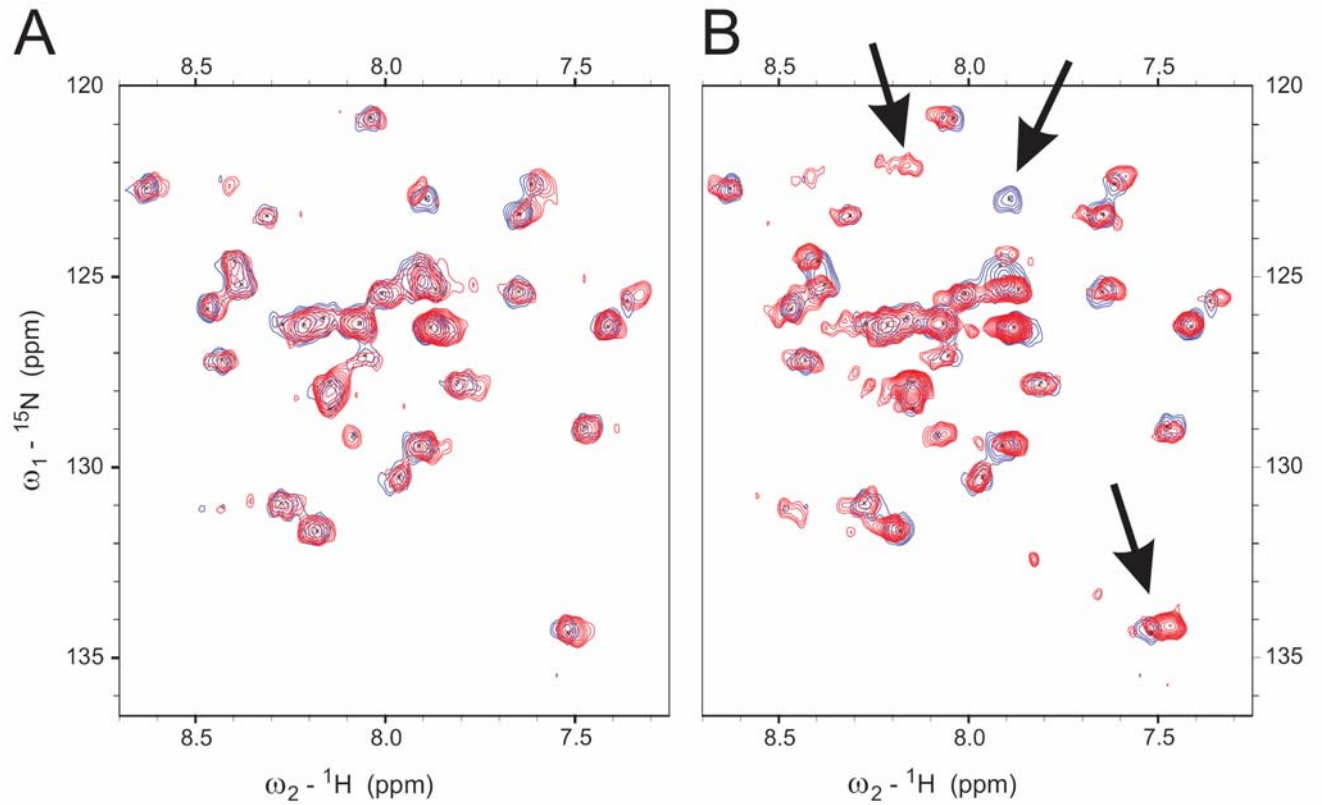


Figure 3.6. Titration of the ^{15}N -labeled pRb with Id-2. The ^{15}N -HSQC spectrum in panel A of the selectively ^{15}N -Leu labeled pRb (construct ABC-CL) mixed in equimolar ratio with Id-2 (R2B) (red); in panel B, pRb mixed in equimolar ratio with the E7-1 peptide (red); pRb alone is shown always in blue. Arrows show peaks that moved.

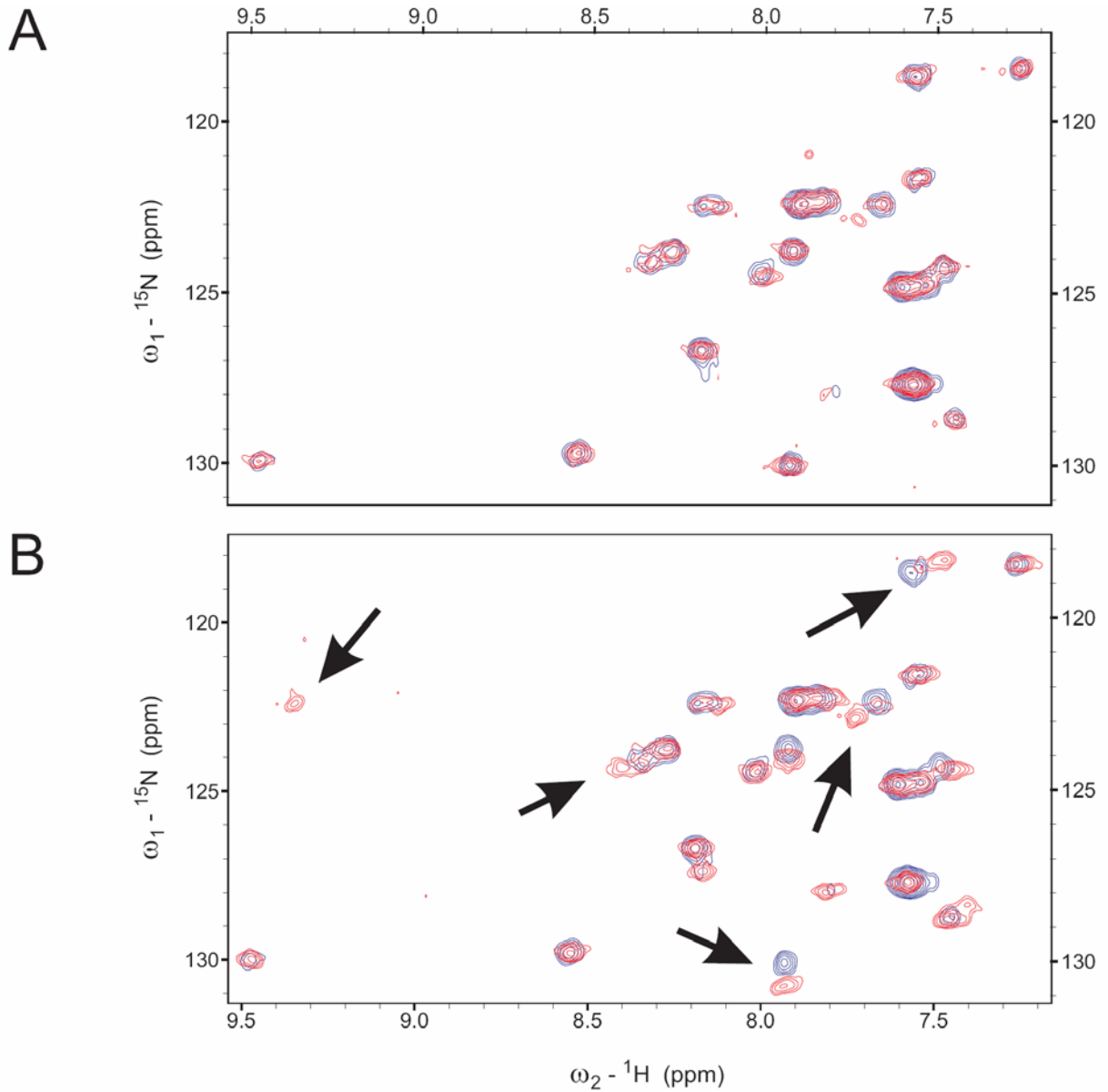


Figure 3.7. Titration of the ^{15}N labeled pRb with MyoD. The ^{15}N -HSQC spectrum in panel A ^{15}N -Lys labeled pRb (construct ABC-CL) mixed in equimolar ratio with MyoD (MCNC) (red); in panel B mixed in equimolar ratio with E7-0 peptide (red); the pRb alone is shown always in blue. Arrows show peaks that moved.

The virus protein HPV16 E7 interacts strongly with pRb. Gel filtration analysis of the 1:1 molar mixture of pRb (ABC-CL) and HPV16 E7 (full length) proteins shows co-elution of both proteins in one peak (Figure 3.8).

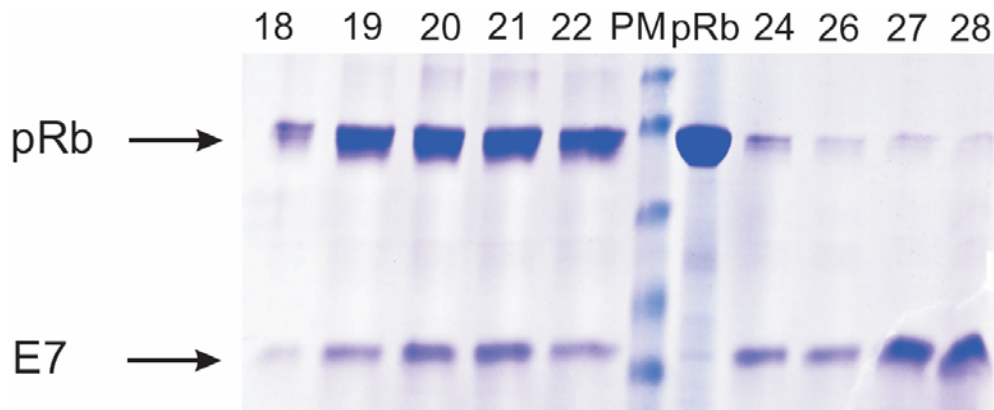


Figure 3.8. SDS-PAGE of consecutive fractions from gel filtration (S-75 Sephadex) analysis of a mixture of pRb and HPV16 E7 (full length) protein. HPV16 E7 and pRb (ABC-C) coelute from the gel filtration column.

3.1.1.2 Discussion

MyoD has two domains that were reported to be responsible for its interaction with the cell cycle proteins. The first is the bHLH domain, which mediates both DNA binding (Ma et al. 1994) and is reported to bind pRb (Gu et al. 1993). The second domain comprises a 15 amino acid stretch claimed to bind to CDK4 (Zhang et al. 1999; Zhang et al. 1999). To check whether these are direct protein–protein interactions we use pull down assay, gel filtration chromatography and NMR spectroscopy. We did not use the yeast two hybrid system or immunoprecipitation as it is known that these methods are prone to give false positive results (Gietz et al. 1997; Toby et al. 2001). We also excluded autoradiographic detection methods for the same reasons (Gu et al. 1993; Lasorella et al. 1996).

We could determine that there is no direct protein–protein interaction between MyoD or Id-2 and pRb *in vitro*. Our results contradict earlier publications (Gu et al. 1993; Lasorella et al. 1996) but are in good agreement with reports which show that MyoD influences cell cycle as a transcription factor (Puri et al. 1997; Cenciarelli et al. 1999; Poleskaya et al. 2001; Magenta et al. 2003). MyoD is widely accepted as a transcription factor that can interact with p300 and CBP proteins of which both are histone acetyltransferases (Halevy et al. 1995). These interactions facilitate binding to a promoter sequence CRE. MyoD together with CBP and CREB proteins trigger transcriptions of pRb (Magenta et al. 2003) and p21 (Halevy et al. 1995; Cenciarelli et al. 1999) – both promoters

of cell differentiation. It was also shown that augmented levels of p57, p27, and probably p21 cause increased stability of MyoD (Zabludoff et al. 1998; Reynaud et al. 1999). These provide a positive feed-back, important at the onset of myoblast differentiation. Without the doubt there is an interplay between myogenesis and the cell cycle. Thus it seems that the primary MyoD function is modulation of genes expression level connected with both myogenesis and the cell cycle.

The genetic interaction between pRb and Id-2 during development using intercrossed pRb and Id-2 mutant mice has been reported (Lasorella et al. 2000). It is possible that Id-2 influences directly or indirectly the expression levels of the pRb gene without physical protein-protein interaction with pRb. It has been shown that the bHLH protein E2A has the ability to block G1 to S phase transition at least through the transcriptional activation of the p21Cip1 gene (Prabhu et al. 1997). Id proteins are able to inhibit E2A mediated expression of the p21Cip1 gene (Prabhu et al. 1997). The p21Cip1 binds to and inactivates CDKs and enhances the activity of pRb, thereby blocking G1 to S phase transition. Therefore, it is possible that Id proteins mediate G1 to S phase transition through the downregulation of p21Cip1 gene expression by heterodimerization with E2A.

3.1.2 *There is no interaction between human or chicken MyoD and CDK6*

3.1.2.1 Results

We tested a possible interaction of the human and chicken MyoD C-terminal 15 aa peptide with CDK6 by a pull down assay, NMR and gel filtration coupled with mass-spectroscopy.

In our pull down assay GST-CDK6 served as a “bait” and the MyoD-C peptide as a “prey”. Amount of eluted MyoD-C did not differ whether GST-CDK6 was present or not. This indicates no binding between these two proteins. Both human and chicken MyoD-C peptides were titrated with CDK6 and no change in peak position was observed in 1D proton NMR spectra (data not shown) indicating that there is no interaction between C-terminal part of MyoD and CDK6.

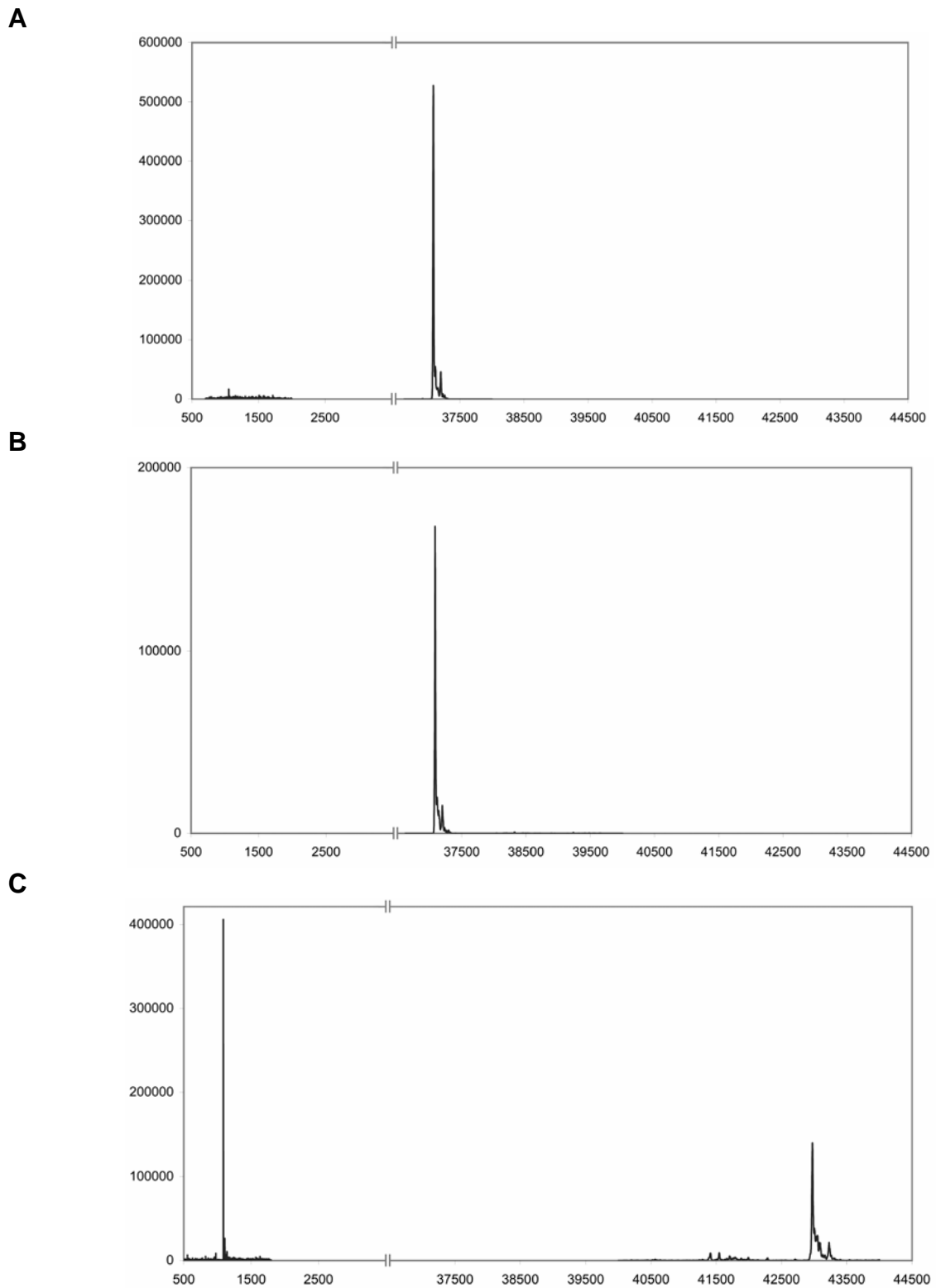


Figure 3.9. Mass spectroscopy results of: (A) gel filtration of the purified sample of CDK6 with the chicken MyoD peptide and (B) human MyoD peptide – the detectable protein corresponds to phosphorylated CDK6; (C) gel filtration purified samples of pRb with the large T antigen peptide – both the protein and the peptide are detectable.

Experiments		Results
	Calculated molecular mass (Da)	Obtained molecular mass from mass-spectroscopy (Da)
CDK6+MyoD	36995.3 CDK6 1769.9 human MyoD 1831.9 chicken MyoD	37084.0 corresponds to CDK6 No mass obtained which corresponds to any of MyoD
pRb+T Ag peptide	42971.4 pRb 1087.1T Ag peptide	42973.0 corresponds to pRb 1087.2 corresponds to T Ag

Table 3.1. Equimolar amount of CDK6 and MyoD peptides were mixed and passed through a S75 gel-filtration column. Fractions corresponding to the complex were pooled, concentrated and checked for the presence of both, CDK6 and the MyoD peptide by mass-spectroscopy. Result showed only one mass corresponding to CDK6 indicating no interaction between CDK6 and any of MyoD peptides. A similar set of experiments was performed for pRb and the T antigen peptide and mass-spectroscopy results showed presence of both species, indicating a physical association between the two. The CDK6 experimental mass higher than calculated is due to of phosphorylation of CDK6 in the Baculovirus expression system.

After gel filtration of an equimolar mixture of CDK6 and the human MyoD–C the peptide fraction corresponding to the CDK6 peak was examined by mass spectroscopy for the presence of the MyoD–C peptide. Mass–spectroscopy showed the presence of only CDK6. The same method showed clearly the presence of T antigen bound to pRb after the sample from gel filtration was analyzed by mass–spectroscopy (Table 3.1., Figure 3.9).

3.1.2.2 Discussion

We showed that there is no direct interaction between the C–terminal of human MyoD (both 15 aa synthetic and the *E. coli* produced peptides spanning residues 212–227 of the human MyoD) or chicken MyoD (15 aa peptide) and human CDK6. Human CDK4 and CDK6 were reported earlier to interact with the chicken C–terminal part of MyoD (Zhang et al. 1999; Zhang et al. 1999). Such cross–species interactions have high degree of uncertainty and do not necessarily represent physiological processes. The amino acids sequence of chicken MyoD that has been mapped to be efficient for binding to CDK4/6 differs substantially from human (Zhang et al. 1999; Zhang et al. 1999). For example, cysteine

residues which could be crucial for the interaction are situated at residues 6 and 13 of human and chicken MyoDs, respectively. The closest homolog for chicken and human MyoDs is *Xenopus-B* and *Pig* MyoDs respectively. CDK4 and CDK6 are highly homologous proteins with sequence homology of 71.1 % (the supplementary materials 1) (Guex et al. 1997; Schwede et al. 2003), which implies high structural homology. CDK4 and 6 have compensating functions in the cell-cycle (Morgan 1997) but there are also some reports suggesting different functions of these two kinases (Ericson et al. 2003). MyoD is proven to enhance expression of CDK's inhibitor p21 (Cenciarelli et al. 1999) which is shown to target and inhibit CDK6-cyclin D complexes in growth-arrested cells.

In summary our results show that there is strong evidence for lack of the direct *in vitro* interaction between two members of the HLH protein family, namely MyoD and Id-2, with pRb. Moreover we have also shown that the C-terminal part of human and chicken MyoD does not interact with human CDK6.

3.2 Photodynamic proteins

3.2.1 *GFPuv*

3.2.1.1 *Results*

We carried out a series of ^1H , ^{15}N , ^{13}C , and ^{19}F NMR studies on the dynamical behavior of GFP. The initial goal of our research was to determine the protonation states of the chromophore that are responsible for the absorption maxima in the optical spectra of GFP. This has proven elusive because the NMR signals of the residues in the chromophore and its vicinity were broadened by conformational motions to such an extent that no assignment of these crucial residues was possible. NMR indicated that the surroundings of the chromophore are structurally mobile, and we studied this phenomenon by making ^{19}F "atomic mutations" of several chromophore protons.

^{19}F NMR characterization of the conformational dynamics was carried out on the cyan fluorescent protein (CFP), a GFP mutant in which tyrosine 66 is replaced with a tryptophan (Seifert et al. 2002). The incorporation of different fluoro-tryptophans (i.e. 4-, 5-, and 6-fluoro-Trp) results in direct post-translational integration of the indole moiety and concomitantly ^{19}F nucleus into the CFP chromophore. Using this approach we were able to

demonstrate that either the CFP chromophore itself or residues in its vicinity interconvert between multiple conformations on the time scale of milliseconds. These slow motions should therefore be considered in the interpretation of structural and spectroscopic properties of the ground states of GFPs. The knowledge of dynamics of CFP opens several possibilities for manipulating the chromophore of CFP in order to improve the fluorescence quantum yield of CFPs (Seifert et al. 2002).

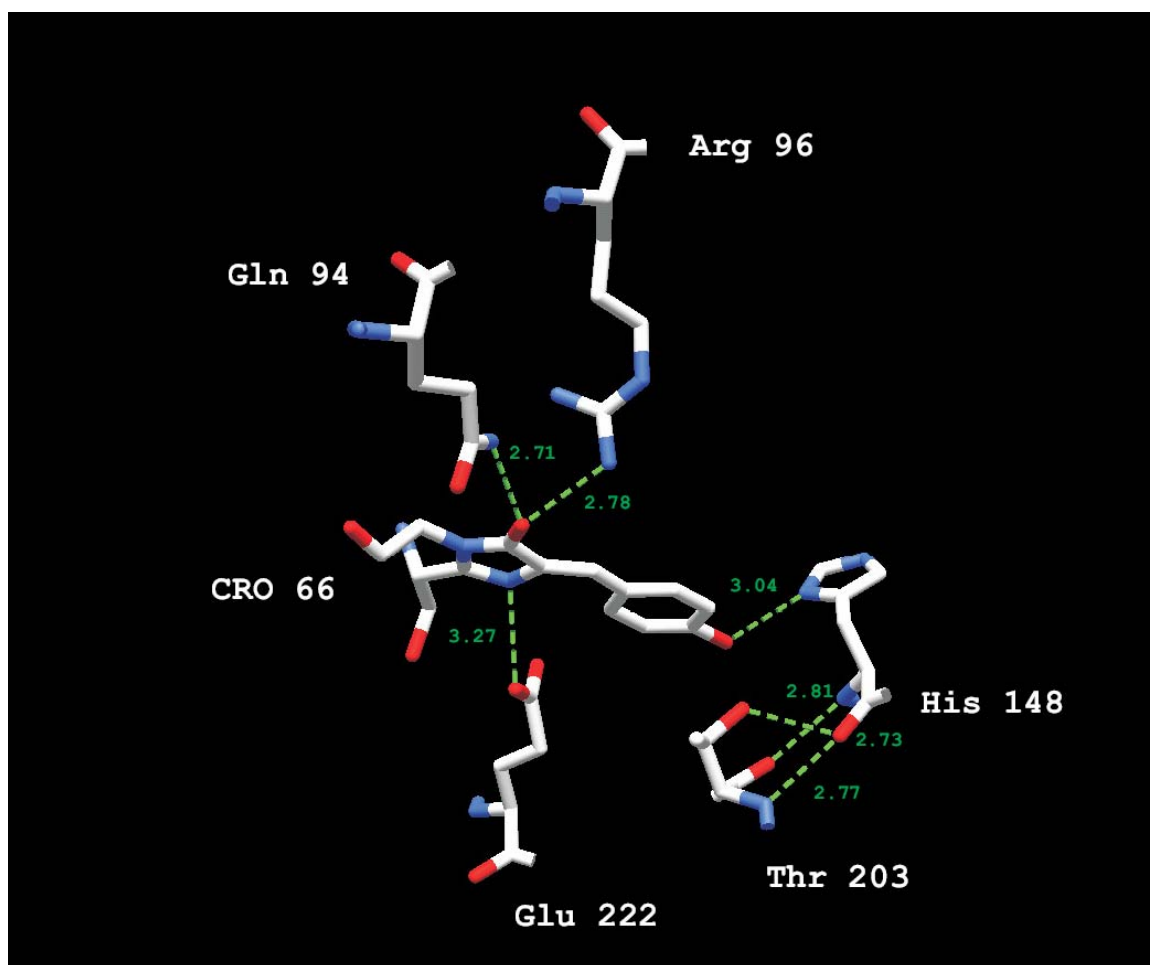


Figure 3.10. Structure of the chromophore of GFPuv. Residues interacting with the chromophore are shown. Green lines represent hydrogen bonds.

3.2.1.1.1 *In vitro* labeling and assignment of residues in the chromophore environment

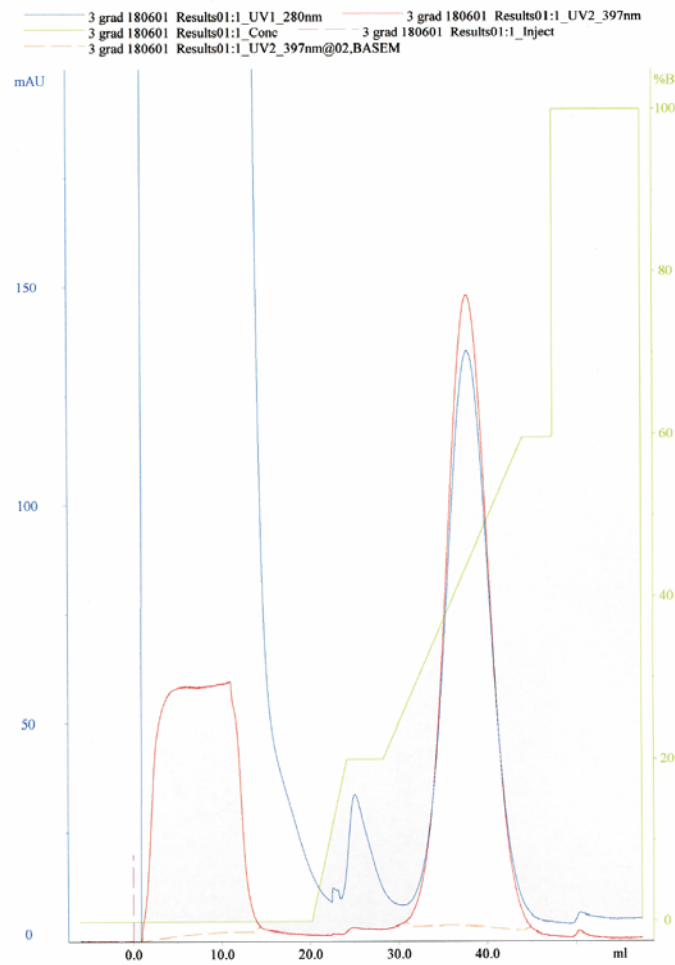
An RTS-500 (description in the Materials and Methods) *in vitro* expression system provides high efficiency and is compatible with T7 expression vectors for *E. Coli* (Hino et al. 2002; Sawasaki et al. 2002; Betton 2003). Thanks to the two compartments (reaction and

supply chambers) the expression reaction could be extensively prolonged compared to classical *in vitro* protein synthesis systems (Sheikh et al. 1990). RTS *in vitro* expression provides advantage for NMR ^{15}N and ^{13}C selective labeling strategies because it does not suffer from the so called “cross-labeling” of amino acids. Cross-labeling is present in all *in vivo* expression systems due to the presence of other metabolic pathways than those involved directly in transcription, translation and energy recycling.

An ^{15}N – histidine selectively labeled sample of GFPuv and its mutant His148 \rightarrow Gly were produced with RTS – 500 (Figure 3.11 A). High purity of samples was monitored as described by Nemetz et al. (2001) (Figure 3.11). We could assign His148 using this sample. His148 is in a direct contact with the protein chromophore. Existence of two peaks corresponding to residue 148 is probably due to switching between two states of the sidechain of His148 as discussed below (Figure 3.12). In this model two conformers of His148 are contributing to two chromophore ground states (Seifert et al. 2003).

Temperature and pH titration, monitored by ^{15}N – HSQC, were carried out on the ^{15}N – His GFPuv sample (Figure 3.13 and 3.14). Only three residues were affected by pH change. They were all situated close to the C-terminus of the protein (His 240) or at flexible loops (His25, His199) (Figure 3.13). This result is in-line with the relaxation data (Seifert et al. 2003). Histidine 25 is additionally situated in close vicinity to Asp21 and Lys26, two residues which we classified as pH sensitive (Seifert et al. 2003). Several more peaks were shown to be temperature dependent (Figure 3.14). Some of them are obviously histidines situated in the C-terminal part (His240) or on the onset of loops (His25, His217, His199). His181 (β -strand 9) (secondary structure elements are shown in the Introduction) is also affected by increased temperature but His81 and His139 remains unaffected. His148 has an interesting behavior. As discussed later we believe His148 is split in the spectrum to two peaks due to two conformational states and the interaction with the chromophore.

A



B

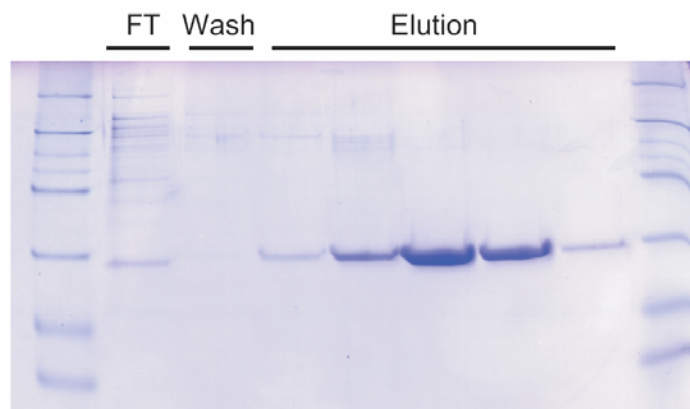


Figure 3.11. (A) Purification of labeled GFPuv was monitored in 280 and 397 nm. This allowed distinguishing between matured GFPuv and *in vitro* impurities – lysate proteins. (B) High purity of the protein was additionally checked by SDS – PAGE.

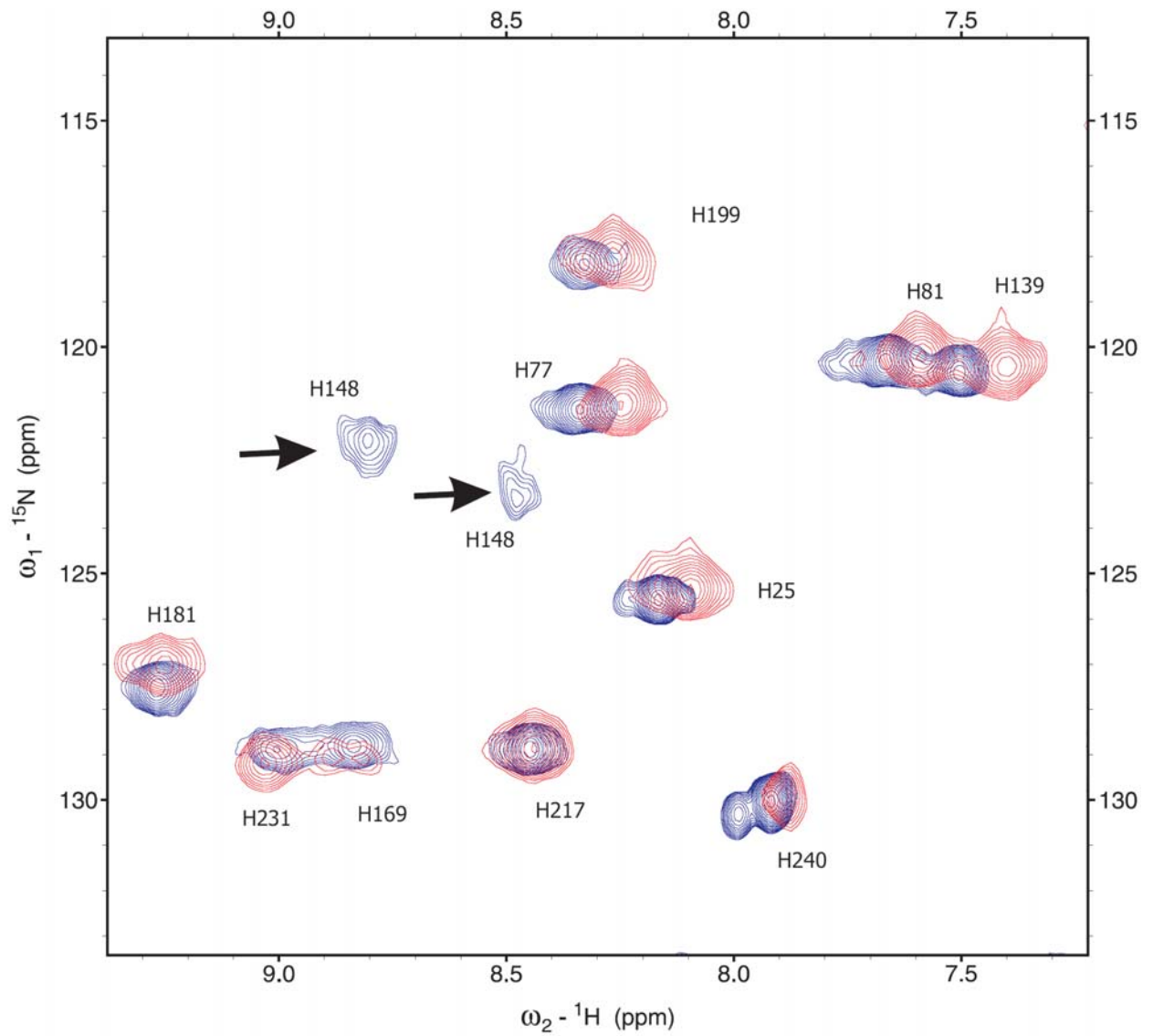


Figure 3.12. ^{15}N – HSQC of the ^{15}N – histidine labeled sample of GFPuv (dark blue) and the mutant His \rightarrow Gly at the position 148 (red). Missing peaks are marked with arrows.

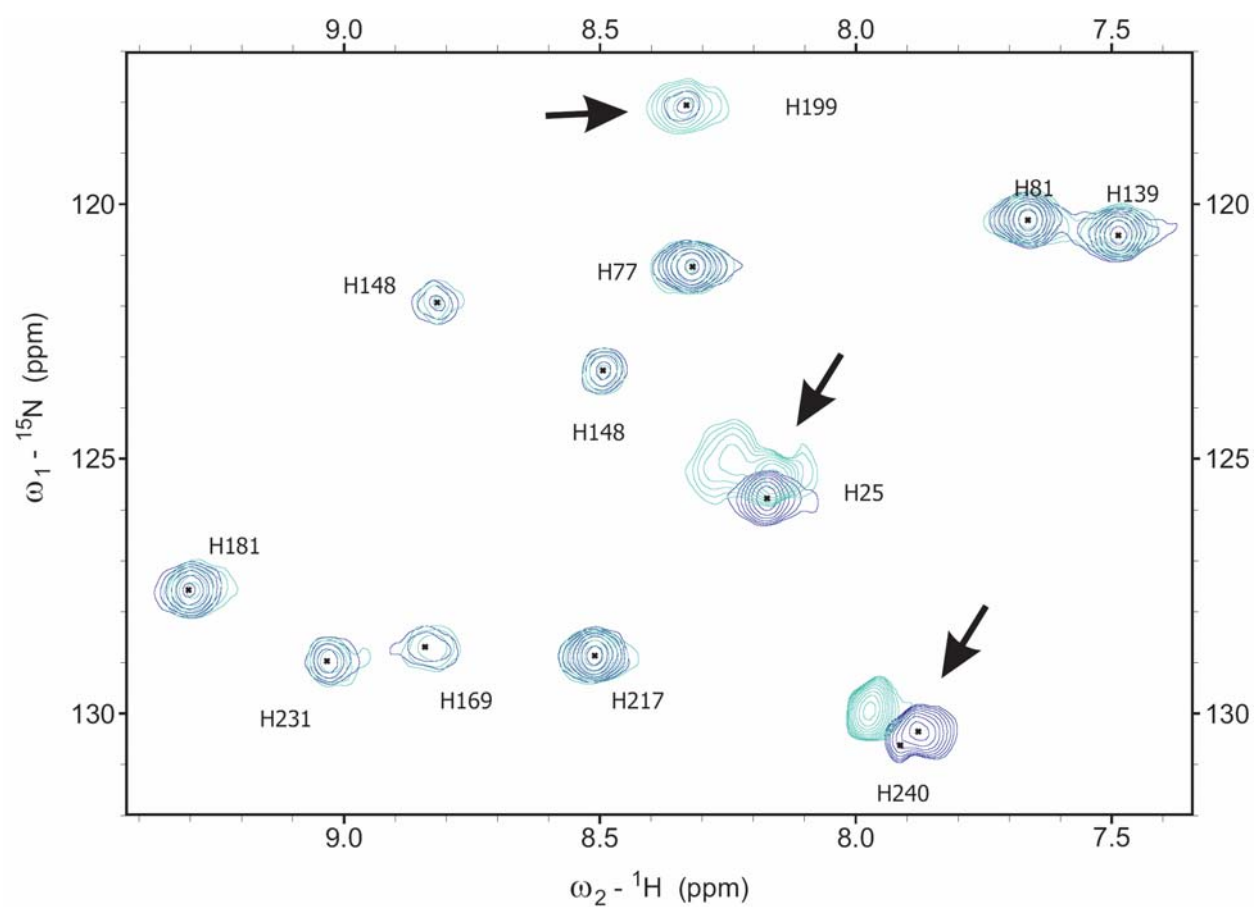


Figure 3.13. The ^{15}N – HSQC of the ^{15}N – histidine labeled sample of GFPuv at two different pH 7,2 (turquoise) and pH 8 (dark blue). Residues which change positions are marked with arrows.

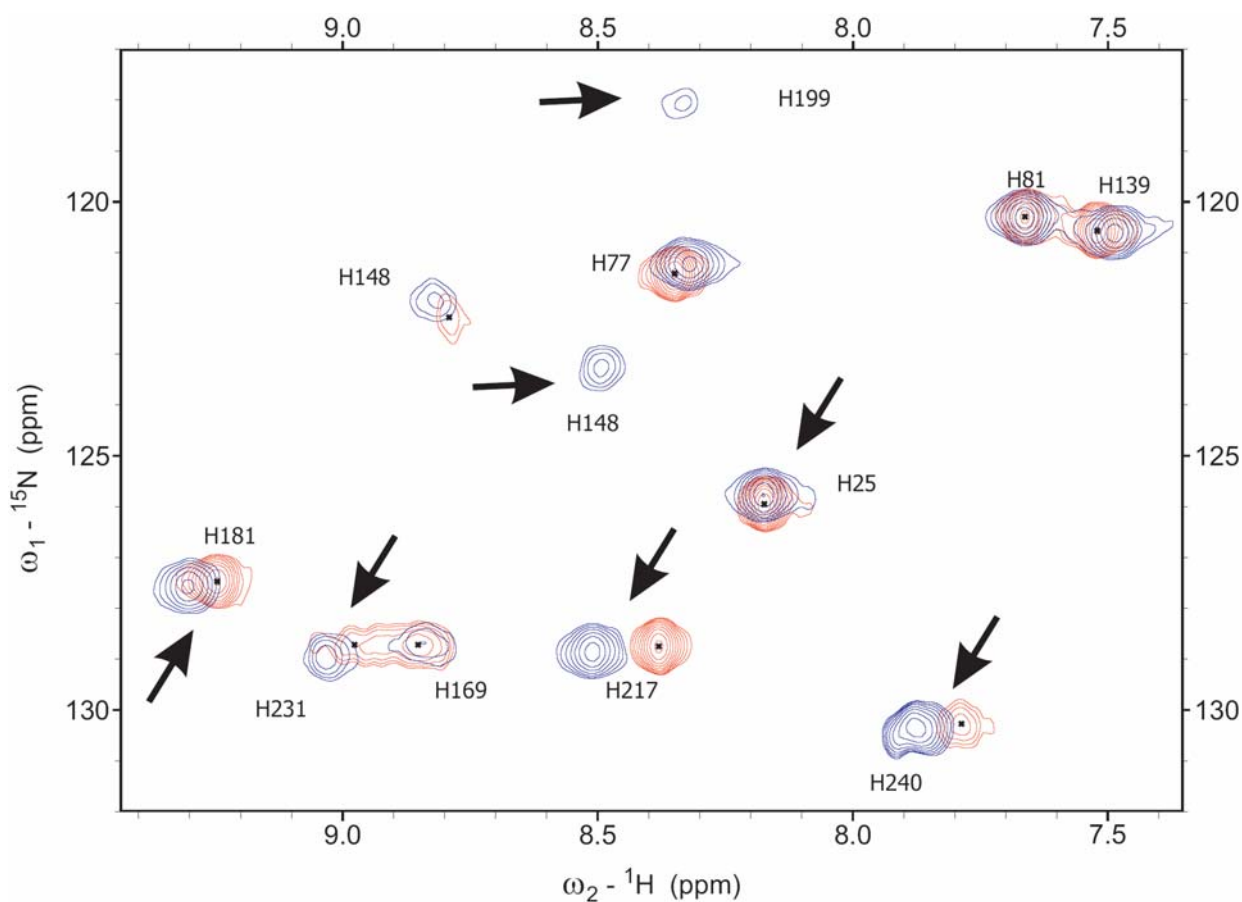


Figure 3.14. The ^{15}N – HSQC of the ^{15}N – histidine labeled sample of GFPuv at two different temperatures, 310K (red) and 290K (dark blue) Residues which change positions are marked with arrows.

3.2.1.1.2 Translational Diffusion

Figure 3.15 A and B illustrate the result of the translational diffusion NMR experiment. Considering the uncertainties of the gradient calibration a diffusion constant of $D = (120 \pm 10) \mu\text{m}^2/\text{s}$ was measured. Using the specific model these results in an apparent mass of GFPuv of 45 ± 10 kDa under the conditions used in NMR experiments. This is significantly larger than 27 kDa of the monomeric GFPuv, but still smaller than 54 kDa expected for a dimer suggesting that GFPuv in solution exists as mixture of monomers and dimers. This is in agreement with the known dimer dissociation constant $K_D = 0.1$ mM of GFPs as determined by analytical ultracentrifugation (Philips 1998). Figure 3.15 C illustrates the GFP dimer interface observed by X-ray crystallography.

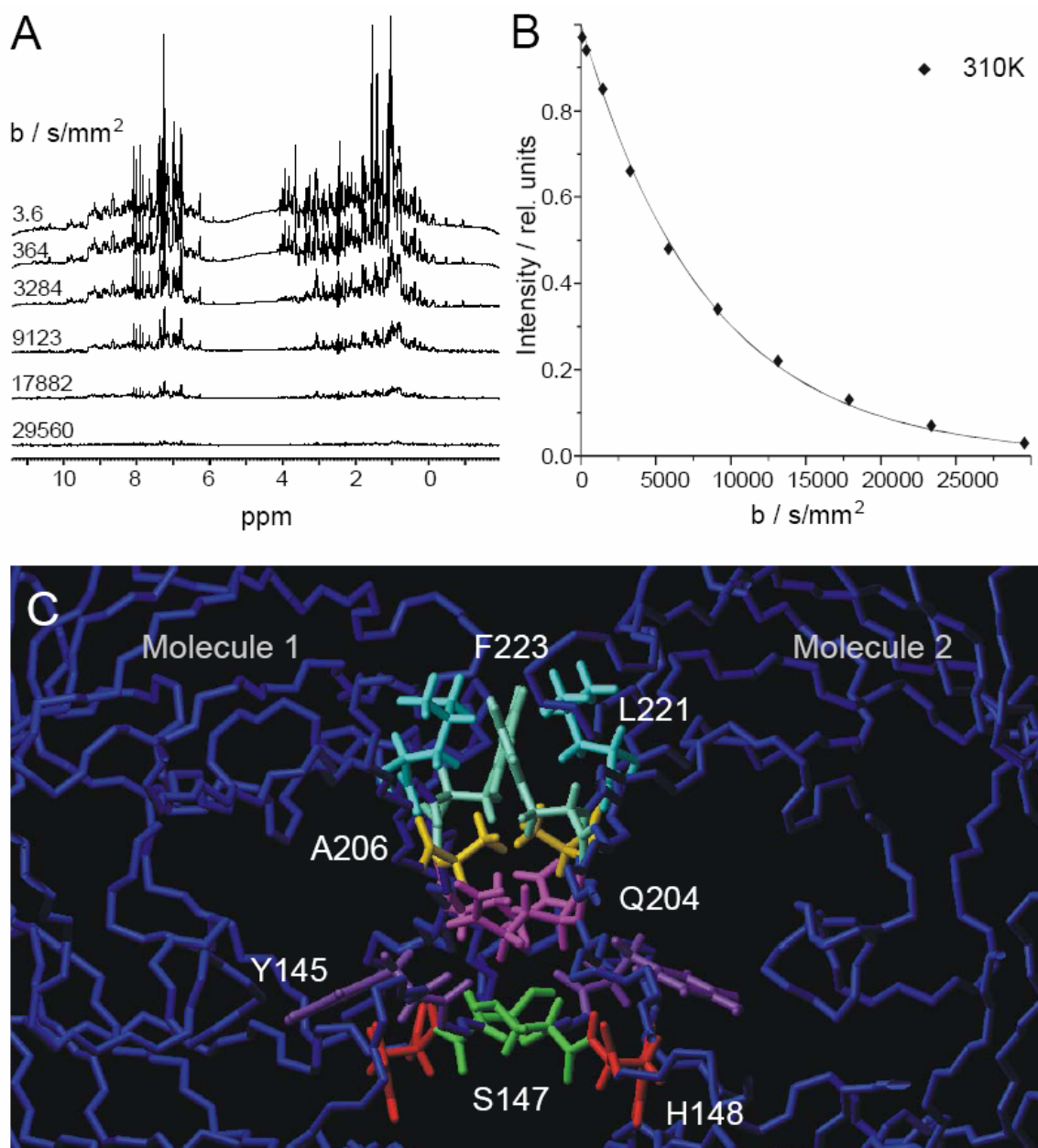


Figure 3.15. NMR measurement of the translational diffusion of GFPuv (A. and B.). C. The dimerization interface of GFP.

3.2.1.1.3 NMR Assignment

Eighty percent of the backbone amide groups of GFPuv were assigned by standard triple resonance methods (Figure 3.16). The assignment was complicated by a substantial amount of missing interresidue connectivities in the triple resonance NMR spectra. The

assignment of the central α -helix and of β -sheets 7, 8, and 10 could not be completed (secondary structure elements declared in Introduction). These difficulties can be explained by the necessity of a neutral pH, the tendency of GFP to aggregate in solution, and the presence of conformational exchange on μ s - ms time scales – as described below – that lead to line broadening.

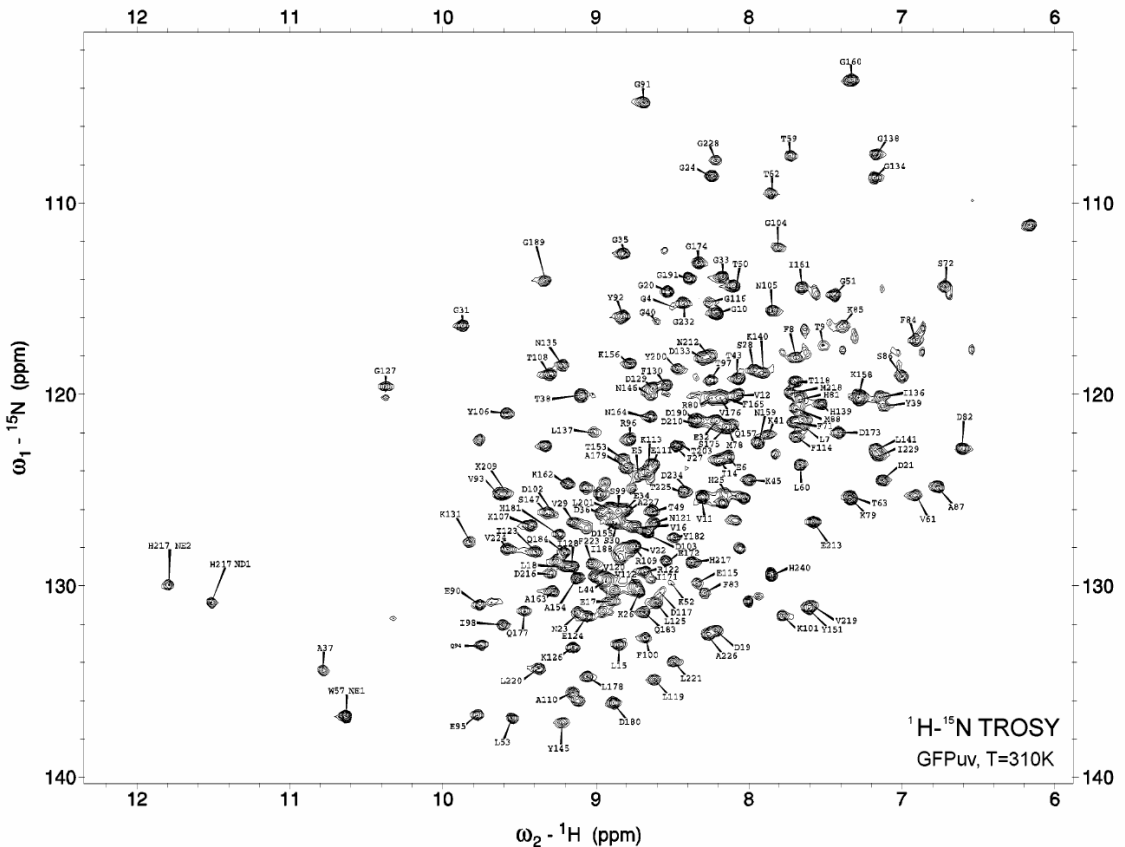


Figure 3.16. A 2D ^1H - ^{15}N NMR TROSY spectrum of GFPuv at 310K. The labels indicate the sequence-specific resonance assignment.

3.2.1.1.4 NMR relaxation

All samples showed essentially the same relaxation behavior. The results of the measurements at 500 and 600 MHz agreed within the expected scaling of the R_1 , R_2 , and hetNOE values with the magnetic field strength. Figure 3.18 summarizes the results of the relaxation measurements on ^{15}N - and ^2H (99%)-labeled GFPuv at 600 MHz. The longitudinal relaxation rate R_1 is 0.5 s^{-1} on average with only the N- and C-terminus

showing higher rates of up to 1.0 s^{-1} . The average transversal relaxation rate R_2 is 27 s^{-1} with lower values mainly for the C-terminal residues. The ^1H - ^{15}N heteronuclear NOE values cluster around 0.8, indicating the absence of pronounced motions on the picosecond to nanosecond time scales for nearly all the residues except the C-terminal ones which exhibit negative NOE values. The ^1H - ^{15}N dipolar- ^{15}N CSA cross-correlation rate η is 15 s^{-1} on average with lower values at the C-terminus of the protein. The overall rotational correlation time of GFPuv (τ_c) was estimated to be 22 ns by calculation of the R_1/R_2 ratios for the backbone nitrogen of residues located in structured parts of the protein ($\text{hetNOE} > 0.6$).

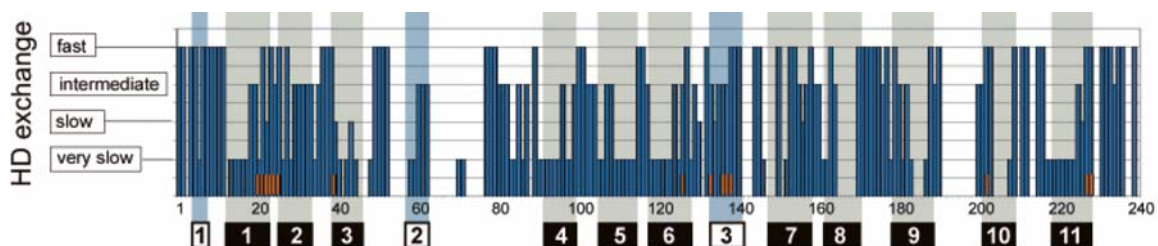


Figure 3.17. H-D exchange rates, and the pH sensitivity of the backbone amide resonances in GFPuv at 310 K. The pH sensitive residues are marked with orange boxes in the diagram of H-D exchange rates. The secondary structure elements of GFPuv are depicted as empty (α -helix) and filled (β -sheet) boxes. Missing bars indicate residues which were not assigned or where no reliable relaxation data could be collected.

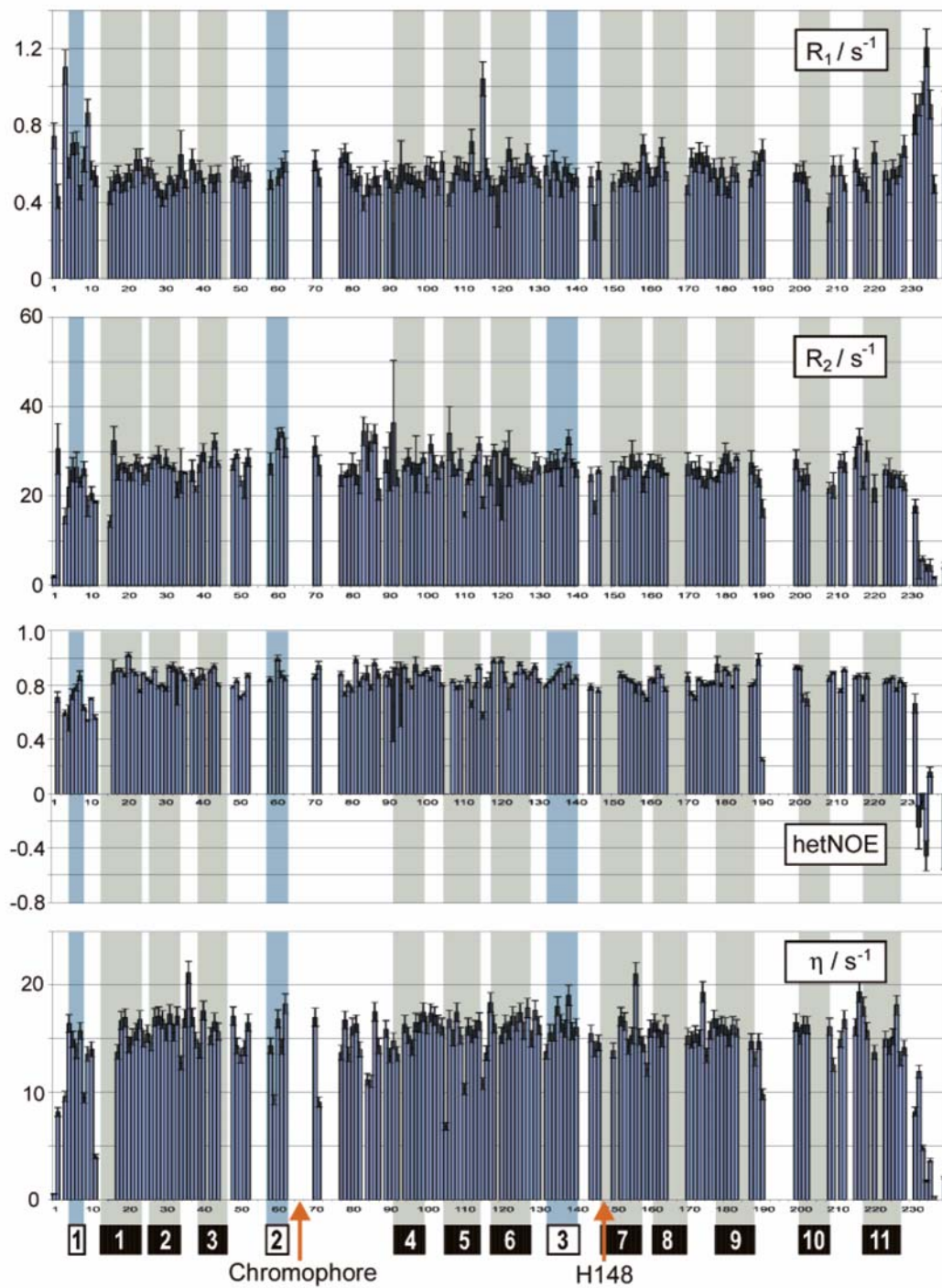


Figure 3.18. Relaxation rates R_1 and R_2 , hetNOE values, and η values as a function of the amino acid sequence of GFPuv. Missing bars indicate residues which were not assigned or where no reliable relaxation data could be collected.

3.2.1.1.5 *pH titration and H – D exchange*

The pH titration of ^{15}N labeled GFPuv revealed that the amide chemical shifts of residues Asp21-Lys26, Gly40, Gly104, Gly116, Gly127, Gly134, Leu137-His139, Thr203, Gly228, and Ile229 are sensitive to pH (Figure 3.17). Except for Thr203, Gly134, and Leu137-His139, all of them are located in loops and exposed to solvent. The region of residues 134-139 corresponds to a solvent-exposed helical section. Residue Thr203 is one of the residues which are in contact with the chromophore. Figure 3.17 shows also the results of the H-D exchange experiment. The resonances were divided into four categories according to the time scale of H-D exchange: fast (completely exchanged after 15 min), intermediate (significant exchange within 12 h), slow (significant exchange within several days and up to 1 month), and very slow (only little exchange after several months). Figure 3.17 shows that the central regions of β -strands 1, 3-6, and 11 exhibit slower average H-D exchange than β -strands 2 and 7-10. The β -barrel of GFPuv is formed by the following sequence: β -strands 1, 2, 3, 11, 10, 7, 8, 9, 4, 5, and 6. The difference in H-D exchange rates in the β -strands provides indications for more pronounced low-frequency motions on one side of the β -barrel corresponding roughly to β -strands 7-10.

3.2.1.1.6 *Mutant His148Gly*

The UV absorption and excitation spectra of GFPuv and GFPuv mutant His148Gly show that the mutation does not dramatically affect the absorption maximum of the chromophore's A state (neutral state) at 398 nm (Figure 3.19). The intensity of peak A in the absorption spectrum and the emission peak at 508 nm is reduced to 82% in the mutant compared to GFPuv. In contrast, peak B (long term anionic state) in the absorption spectrum is reduced to 26% in the mutant protein. Peaks A and B are usually assigned to the neutral and anionic state of the chromophore, respectively. The emission at 510 nm is reduced to 80% in the mutant protein, which is agreement with the reduced absorption of state A.

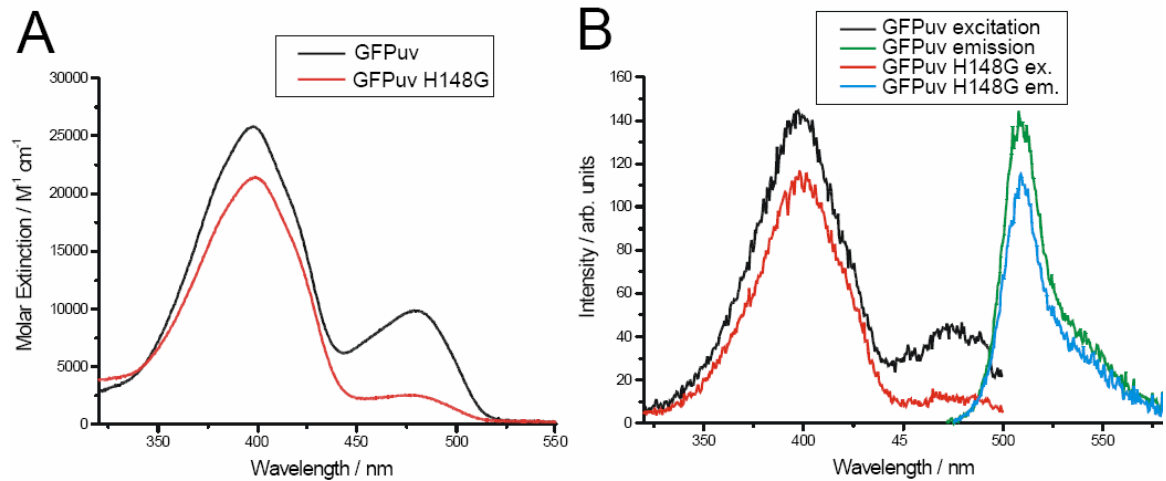


Figure 3.19. A. UV/VIS absorption spectra B. Excitation and emission spectra of GFPuv and GFPuv mutant H148G.

Mutation His148Gly of GFP leads to the appearance of double peaks for several residues in the HSQC spectrum of GFPuv His148Gly (Figure 3.20).

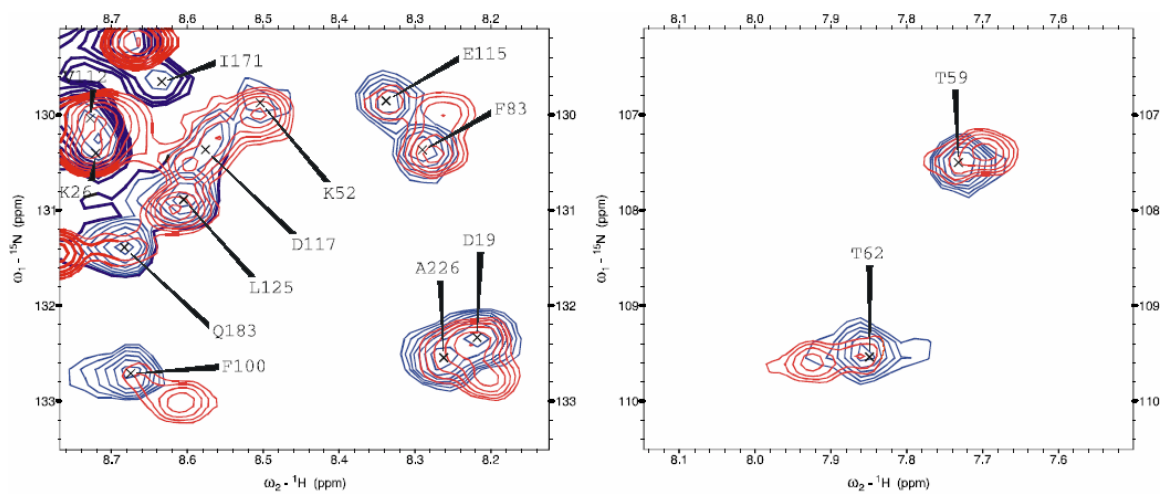


Figure 3.20. Two regions from the ¹H-¹⁵N HSQC spectrum of GFPuv (blue) and GFPuv His148Gly (red) are shown exemplifying the appearance of double peaks in the mutant His148Gly for several backbone amide groups.

The peak volume ratio of minor and major peaks was on average 0.45:1 at 310K. Minor peaks corresponded to the original resonances in the spectrum of GFPuv. The small difference in chemical shift of major and minor peaks of approximately 30 - 60 Hz indicates a timescale of exchange at the order of 0.01 s to 0.1 s. At a temperature of 290° K the minor peaks seem to be even more pronounced although broadening of the resonances due to a higher solvent viscosity complicated the detection of several resonances. These temperature dependent double peaks provide an indication for a slow exchange processes affecting the backbone structure of the mutant. The following residues with known assignment are affected by these exchange processes: Thr43, Thr49, Thr59, Thr62, Ser72, Phe83, Thr97, Ile98, Phe100, Phe130, Lys158, Leu178, Tyr182, Gln183, Thr203, His217, Val224 and Ala226. In addition the heteronuclear NOE values of the minor confirmation are significantly reduced for the following residues (data not shown): Gly33, Glu34, Gly40, Lys41, Phe84, Phe100, Gly116, Asn164 and Phe165. The heteronuclear NOE values of the major conformation are comparable to those obtained for GFPuv. Figure 3.20 maps the location of residues that were assigned and unambiguously affected by exchange on the crystal structure of GFP.

These residues cluster in the central α -helix (as far as assigned) and β -strand 4, however, the effect is not limited to residues near His148 and many residues distributed all over the β -barrel are affected.

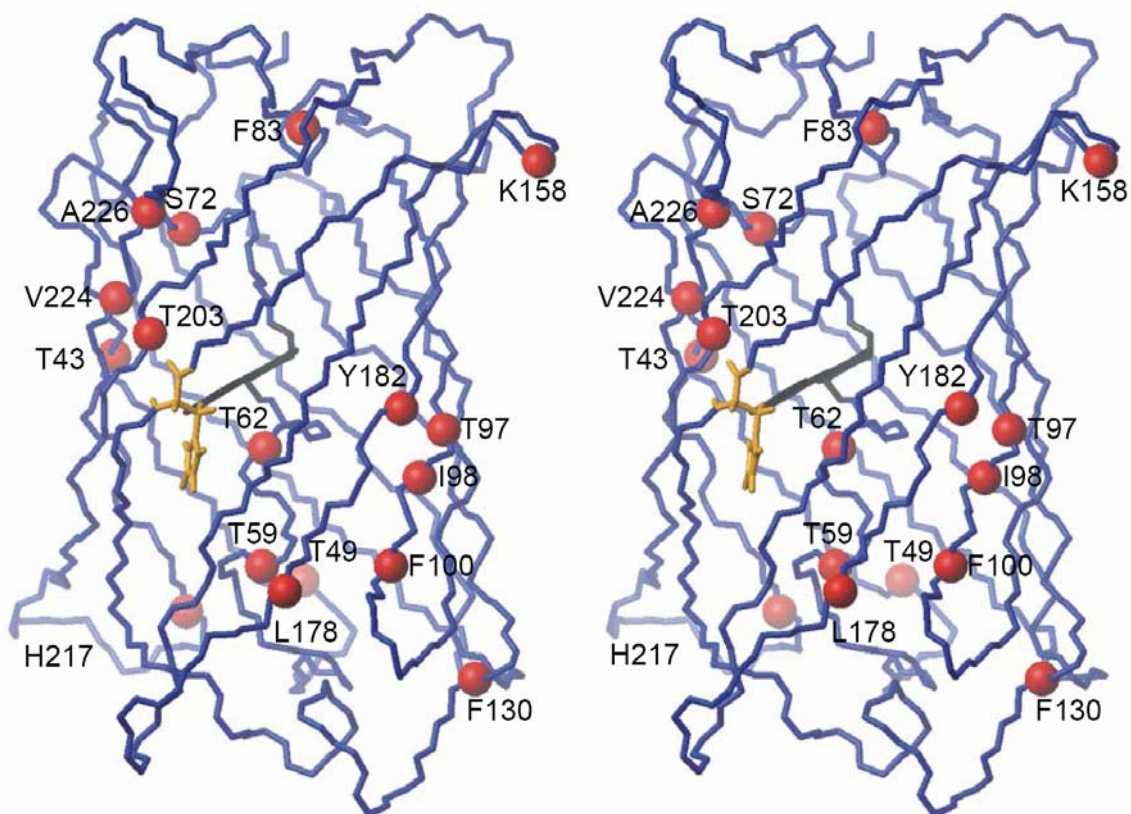


Figure 3.21. Dual conformation in GFPuv His 148 Gly. The affected amide nitrogens are shown as red balls with respect to the GFP structure. The chromophore and H148 are depicted in gray and orange, respectively.

3.2.1.2 Discussion

It is usually assumed that protein folding is driven by the thermodynamic advantage inherent in the sequestration of hydrophobic residues in the interior of the protein away from the solvent (Prendergast 1999; Dobson 2003). However, in the core of GFP, i.e. around the chromophore, there is a surprising number of polar residues (Tsien 1998). Therefore other nucleation points for protein folding are likely to exist. Hydrophobic residues Phe27 and Leu60 are conserved in all proteins mentioned here suggesting that these residues might be important for the stable protein structure, presumably during the folding process. The conserved amino acids Phe27 and Leu60 together with Leu125 form a hydrophobic core within the barrel that may play a role in positioning the central α -helix during protein folding. In addition, conserved glycine residues may also play an essential role for the GFP folding. It has been shown that the Thr22Gly mutation in *Drosophila* drk protein, which

restores the otherwise highly conserved glycine residue in the diverging β -turn connecting β -strand 3 and 4 of SH3 domains, significantly stabilizes the drk SH3 structure (Mok et al. 2001). It was proposed that glycines might serve to be hinges enabling the movement of loops. Such hinges may also be important for the GFP β -barrel. The apparent protein mass of GFPuv of approximately 45 kDa determined from translational diffusion indicates the presence of a mixture of monomers (27 kDa) and dimers (54 kDa) in solution. The extent of self-association in the GFPuv sample is obvious when comparing the overall rotational correlation of GFP (238 residues, $\tau_c = 22$ ns) to the one of e.g. maltose binding protein (MBP (Gardner et al., 1998), 370 residues, $\tau_c = 16$ ns), which was measured at the same temperature as GFP. It was proposed that the mutations present in GFPuv, Phe99Ser, Met153Thr, and Val163Ala, reduce the aggregation tendency (Fukuda et al. 2000). Our results clearly show that these mutations cannot prevent self-association of GFPuv under the conditions needed for NMR spectroscopy, i.e. the high protein concentration of approximately 1 mM. Additionally, under these conditions one-dimensional ^1H NMR spectra (Figure 3.22) of GFPuv and Clontech EGFP (Phe64Leu, Ser65Thr) show no significant difference in ^1H line width between GFPuv and EGFP.

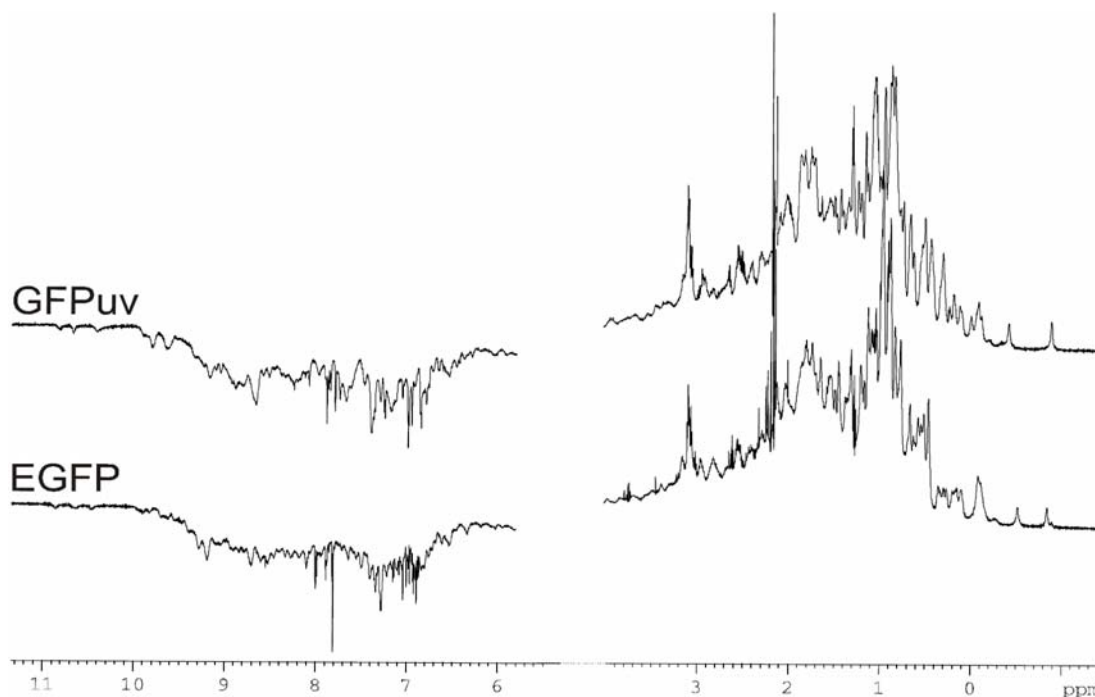


Figure 3.22. ^1H NMR spectra of GFPuv and EGFP (Clontech) recorded at of 310 K and a proton frequency of 600 MHz.

Also mutations like Ala206Lys, Leu221Lys, and Phe223Arg (Zacharias et al. 2002) are known to reduce hydrophobic interactions and might therefore allow for NMR samples with less aggregation. It is not yet clear if these hydrophobic interactions play a physiological role for any kind of membrane association or interaction with the light emitting protein aequorin. It is known that other GFPs, like Renilla GFP, are obligate dimers (Tsien 1998) whereas DsRed forms a tetramer (Campbell et al. 2002; Zacharias et al. 2002). In general, the self-association of GFPs is a relatively conserved feature. Recently it has been speculated that a relationship between oligomerization and habitat temperature may exist (Campbell et al. 2002; Zacharias et al. 2002). Hydrophobic interactions may also be relevant for the interaction of GFP and the light-emitting protein aequorin. Remarkably, the structurally very similar (2.6 Å rmsd compared to GFP) nidogen-1 is supposed to play a role in membrane assembly (Hopf et al. 2001). Figure 3.23 maps experimental findings on the GFP secondary structure. With 80% of the backbone amides being assigned, some important conclusions can be drawn from the relaxation measurements (Figure 3.18): (a) the overall structure of GFP is rigid on the picosecond to nanosecond time scale. This is in agreement with molecular dynamics simulation, which predicted a RMS deviation of only 0.9 Å during a 1 ns period (Helms et al. 1999). (b) A significant variation of the η/R_2 values can be observed in β -sheets 4, 5, and 6 and especially in the first α -helix where the η/R_2 values are lower at the beginning and at the end of the sheets compared to the center of the sheets. This indicates motions on the μs – ms timescale in the turns connecting the sheets. (c) One side of the protein, corresponding to β -strands 7, 8, 9, and 10, shows increased HD exchange rates. Increased HD exchange rates in β -sheets are usually interpreted as the result of hydrogen bonds being broken by conformational fluctuations on the μs - ms time scale (Dempsey 2001).

Murzin et al. 1994) is fulfilled to good approximation. Therefore inefficient side chain packing is unlikely to be the origin of the disturbed hydrogen-bonding network.

Several indications underline the importance of dynamical processes around β -strand 7: Firstly, according to molecular dynamics simulations conformational fluctuations occur in the cleft between β -sheets 7 and 8, mainly connected to flipping of the Arg168 sidechain (Helms et al. 1999) at least on a nanosecond timescale. Secondly, ^{19}F NMR studies on the cyan variant of GFP showed that His148 is likely to be involved in a slow exchange process (Seifert et al. 2002). Finally, it has been speculated that conformational fluctuations are the origin of the pH sensitivity of the GFP chromophore (Haupts et al. 1998). For example, mutations at the positions 147, 149, 164-168, 202, and 220 (located in β -strands 7, 8, 10, and 11) have been shown to increase pH-sensitivity of the GFP chromophore (Miesenbock et al. 1998). This effect is supposed to be caused by an increased conformational flexibility around the chromophore. Therefore a dynamic process, e.g. an exchange between different backbone conformations, is most likely the origin of the increased hydrogen exchange rates, which may also explain the difficult resonance assignment in this part of the protein.

In summary, the observed H – D rates indicate a higher degree of conformational fluctuations that disturb the inter-strand hydrogen-bonding network around β -strand 7. NMR methods exist to measure hydrogen bonds directly, e.g. via the $^1\text{J}_{\text{NC}}$ coupling constants, but the high molecular weight and aggregation tendency of GFP precludes the practical usage of such sequences (Juranic et al. 1996). From the point of view of fluorescence the observed conformational flexibility is not optimal for a high quantum yield. For example, such motions may allow fluorescence quenchers like oxygen to enter the β -barrel. In addition, the backbone flexibility may also affect the chromophore-stabilizing network and therefore increase the rate of photoisomerization. To further optimize fluorescence properties of GFPs by reducing conformational flexibility, clarification is needed whether this conformational flexibility is necessary for protein folding or if suitable mutations can reduce such structural instabilities. However, since the β -strand 7 is interrupted around residue 148 in most GFP structures and surprisingly also in the nidogen-1 structure, where Tyr540 corresponds to His148, the presence of a predetermined breaking point at this point is suggested. It has been speculated that the closure of the cleft between β -strands 7 and 8 is one of the last events in the formation of the β -barrel upon folding of the protein (Helms et al. 1999). It is therefore more likely that the mentioned flexibility is an intrinsic property of the GFP fold that might result from its folding. To investigate the relationship between structural properties of β -

strands 7, 8, and 10, and GFP function (i.e. fluorescence), histidine 148 located in β -sheet 7 was mutated to glycine. The most apparent difference in the optical properties of GFPuv His148Gly compared to native GFPuv is the strong reduction of absorption at 478 nm, referred as band B arising from the anionic state of the chromophore. In addition, results here show that this mutation leads to the appearance of an allosteric slow exchange processes in the protein backbone. Therefore this mutation destabilizes the overall structure of the protein, leading either to an increase in the amplitude of low-frequency motions or a shift in time scale which makes the motions visible in the HSQC spectrum. The effect observed in GFPuv His148Gly has to clearly distinguish from the slow exchange process observed in CFP (Seifert et al. 2002), which takes place on a much shorter timescale (1.3 ms) compared to GFPuv His148Gly ($k_{\text{ex}} \approx 0.01 - 0.1$ s). The most interesting question that arises from the discussed observations is whether a mechanistic link can be drawn from structural dynamics derived from NMR data to optical properties. First of all, since light absorption takes place on the femtosecond time scale – much faster than any structural process in proteins – UV absorption spectra display only an instantaneous average of the immediate chromophore environment (Lakowicz 1999). Secondly, for the interpretation of fluorescence data it is always necessary to infer the physiochemical behavior of the environment from measured photophysical parameters on the assumption that the photophysical responses are understood a priori, i.e. a plausible model of the chromophore and its environment is available (Moncrieffe et al. 2000). For GFP a model of photophysics has been derived from crystal structures in combination with time-resolved fluorescence measurements (Tsien 1998; Prendergast 1999). In the case of GFPuv His148Gly the main question is whether the observed strong reduction in absorption of the anionic state B – whereas the fluorescence emission is nearly unchanged – is related to the structural inhomogeneity as seen in NMR spectra. The creation of the non-equilibrium excited I^* state – which is supposed to give rise to the fluorescence emission – by deprotonation of the chromophore's phenolic oxygen seems to be quite unaffected by the mutation of His148Gly. However, the anionic state's $B \rightarrow B^*$ transition is significantly less prominent which may be attributed to a shifted ratio of neutral to anionic states in the chromophore's thermodynamic equilibrium. So, it seems that the substitution of His148 with glycine and thereby removing a major part of the internal hydrogen bonding network affects only the state of chromophore protonation in equilibrium. The access of solvent molecules through the structural

fluctuations in the β -barrel alters the immediate chromophore environment but still allows a fast and efficient creation of the fluorescent state by chromophore deprotonation.

3.2.2 *¹⁹F fluoro tryptophane GFP*

3.2.2.1 *Results*

3.2.2.1.1 *UV and Fluorescence Spectroscopy*

As expected, the UV absorption profile as well as fluorescence spectra of EGFP (F64L, S65T) were not affected by fluorination of W57, which is placed about 15 Å apart from the chromophore. In fact, the single absorption peak at 488 nm of both EGFP variants as well as the emission maximum at 510 nm clearly indicate the anionic state of the chromophore (Figure 3.24 A). Conversely, native ECFP (F64L, S65T, Y66W, N146I, M153T, V163A) exhibits two chromophore absorption bands (434 and 452 nm) and two distinct fluorescence emission maxima (476 and 500 nm). Fluorination of W57 and the chromophore indole induced 4 nm blue shift in the absorption maxima of 6FW-ECFP (430, 448 nm) as well as a decrease in the absorbance intensity by about 10 %. These effects are even more pronounced in fluorescence properties of 6FW-ECFP: while emission maxima are negligibly blue shifted by about 2 nm, the fluorescence intensity is lowered by about 50 % compared to the native ECFP (Figure 3.24 B). It should be however noted that basic features of spectral behavior of ECFP (i.e. the presence of two spectral bands in UV and fluorescence) are not changed upon fluorination. A detailed study on the effects of fluorination on the optical properties of GFP variants has been reported by Budisa et al. (2002).

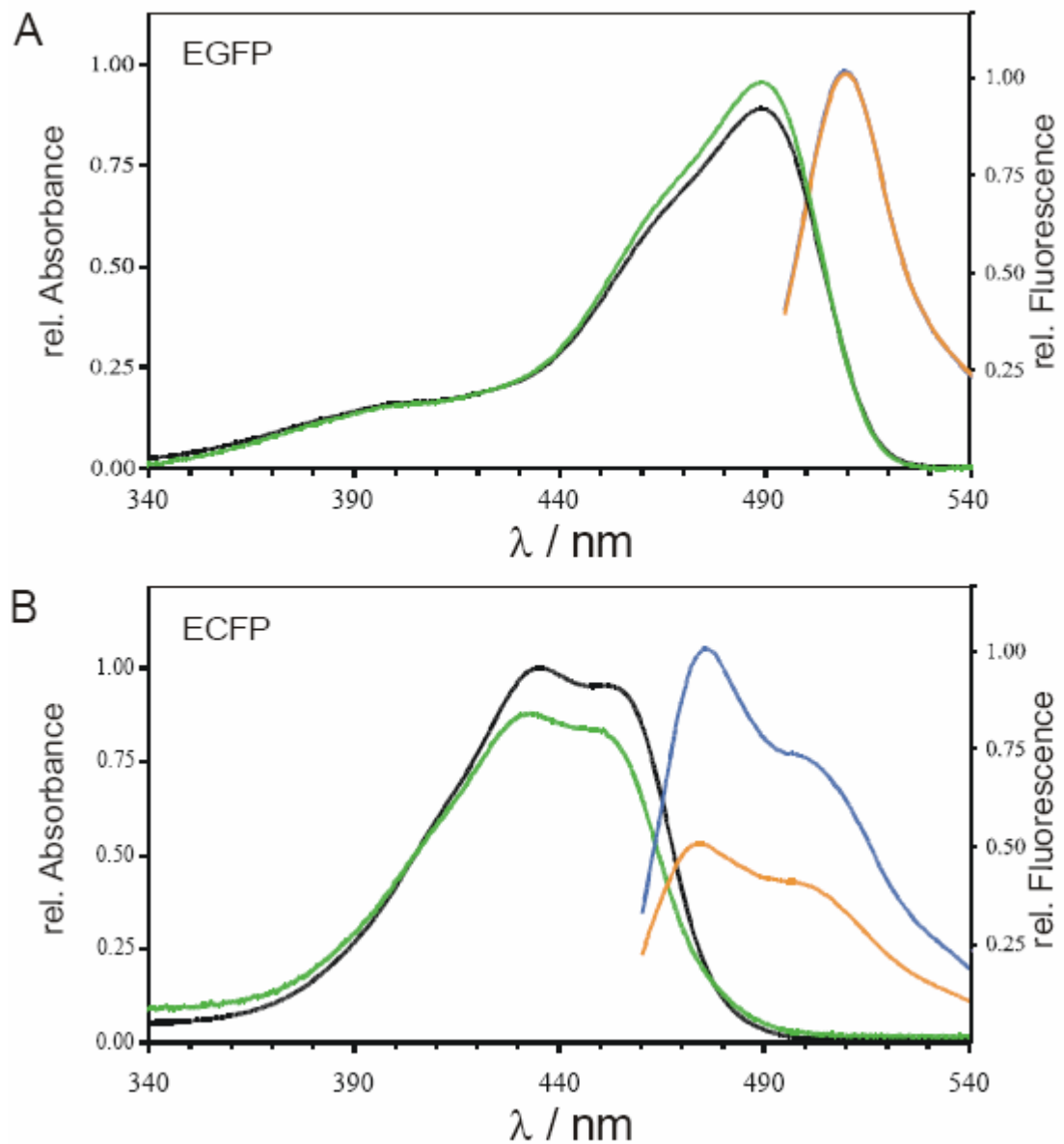


Figure 3.24. UV absorption and fluorescence emission spectra of EGFP at neutral pH (A. black and blue), 6FW EGFP (A, green and orange), ECFP (B. black and blue) and 6FW ECFP (B, green and orange). Note that the absorption band around 395 nm characteristic for wild type GFP is suppressed in EGFP and its fluorinated counterpart. Only one absorption band at 488 nm and emission maxima at 510 nm are characteristic for the anionic state of the chromophore in EGFP. On the other hand the UV-absorbance and fluorescence spectra for both native and fluorinated ECFP are characterized by two spectroscopically distinguishable states.

3.2.2.1.2 ^{19}F NMR spectroscopy of fluoro tryptophane EGFP, ECFP, EYFP

All ^{19}F NMR spectra were referenced to an external reference consisting of 10 mM TFA in 90% H_2O , 10% D_2O at a temperature of 303 K. For a typical one-dimensional ^{19}F NMR spectrum of 14 kHz sweepwidth up to 16k scans with an interscan delay of 1 s were recorded. Baseline distortions caused by a relatively long prescan delay were corrected by manual baseline fitting. The ^1H - ^{19}F J-coupling constants in 4-, 5- and 6-fluoro-tryptophan were measured in ^{19}F NMR spectra of pure tryptophan samples. To evaluate the effect of J-coupling on the results of lineshape fitting experimentally, spectra of 6FW CFP with and without proton decoupling were recorded. Despite of a strong reduction of signal intensity upon decoupling no difference in lineshape of the W66 resonances was visible (data not shown). Therefore it was concluded that the calculated values for the rate constants are not affected by J-coupling. pH titrations of the 4FW and 6FW labeled CFP sample were carried out by dissolving the protein in 50 mM borate buffer at pH 9.1 with 100 mM NaCl and then lowering the pH by adding the appropriate amount of 1% H_3PO_4 for each step. ^{19}F spectra at pH 9.1, 8.1 and 7.1 were recorded. In order to record a ^{19}F NMR spectrum of unfolded 6FW CFP, urea was added and heating to 95°C for 10 minutes was applied. To analyze the influence of protein concentration on ^{19}F NMR spectra a sample of 6FW ECFP was diluted in 4 steps to one sixteenth of the original concentration. The ^{19}F NMR spectra of the reference and of the diluted samples were recorded with 8k scans (reference, 8 mg/ml, 0.3 mM monomeric), 16k scans (4 mg/ml, 0.16 mM), 32k scans (2 mg/ml, 0.08 mM), 64k scans (1 mg/ml, 0.04 mM) and 128k scans (0.5 mg/ml, 0.02 mM). The effect of UV irradiation on 6FW ECFP was checked by recording ^{19}F NMR spectra after irradiating the NMR sample with a 450 W Xe lamp (~ 5 mW light power at 400 nm) for 4 hours.

The ^{19}F NMR resonance of free 4-fluoro-tryptophan is split into a quartet ($^3\text{JFH} = 12$ Hz, $^5\text{JFH} = 4.6$ Hz), whereas 5- and 6-fluoro-tryptophan show a sextet structure ($^3\text{JFH} = 11$ Hz, $^5\text{JFH} = 5.5$ Hz) (data not shown). These splittings are obscured in the broad protein signals when these tryptophans are incorporated into EGFP and ECFP proteins. The ^{19}F NMR spectra of ECFP and EGFP labeled with 4-, 5-, or 6-fluoro-tryptophan exhibited four and two resonances, respectively (Figure 3.25). CFP contains only two tryptophan residues, W57 and W66, whereas GFP has only one tryptophan, W57. This clearly indicates that each of the two tryptophans in CFP exists in two states. The more and less populated states of W57 and W66 are marked with subscript A and B, respectively (Figure 3.25). Since one of the CFP tryptophans, W66, is incorporated into the CFP chromophore during protein folding

by cyclization of the backbone, two states exist within the chromophore. The resonances of 4FW and 6FW ECFP can be unambiguously assigned by comparison with the EGFP spectra. The assignment of the resonances of 5FW ECFP is complicated by the fact that a second peak is not visible in the 5FW EGFP spectrum. Since W66 and W57 are incorporated in the protein in a 1:1 ratio, the sum of integrals for both states of each residue must give the same value. The ratio of the integrals of states A and B for W57 and W66 are consistent within the 4FW and 5FW CFP samples ($W66A/W66B = 0.3$ and $W57A/W57B = 0.2$). 6FW CFP shows integral ratios of $W66A/W66B = 0.6$ and $W57A/W57B = 0.6$. The difference in chemical shift for the two states of W66 is 2.6 ppm in 4FW and 6FW ECFP. The ^{19}F NMR spectra of 4FW and 6FW GFPuv did not show double peaks (data not shown) indicating a different timescale of exchange.

The appearance of two peaks for each tryptophan in the spectra of ECFP and EGFP and the different broadening of those peaks provided a first indication for the existence of slow exchange processes between different states of the tryptophan sidechains. This is especially important since it gives a strong hint that a slow exchange process exists within the chromophore of CFP.

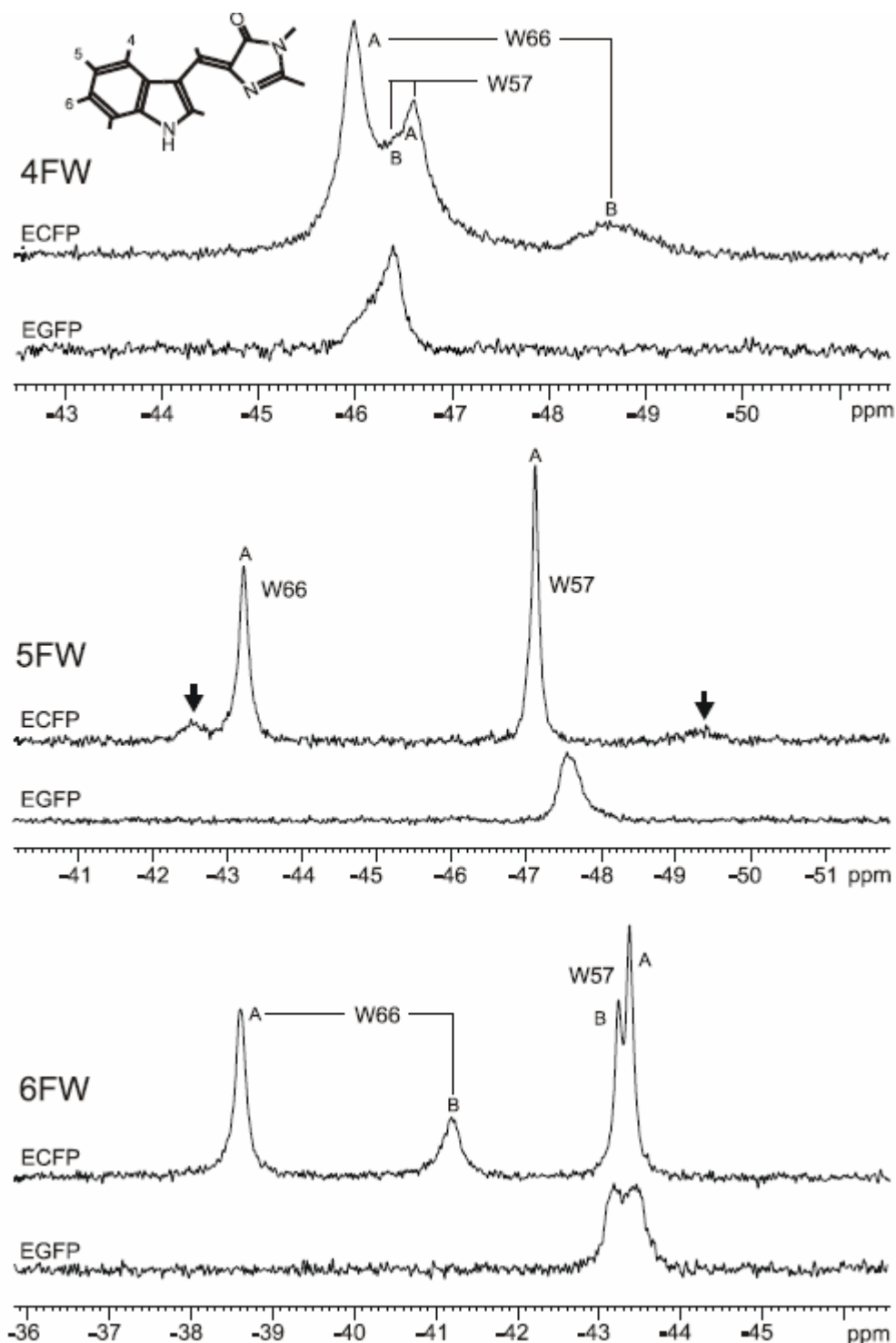


Figure 3.25. ^{19}F NMR spectra of ECFP and EGFP labeled with 4FW, 5FW and 6FW: The chromophore of CFP and the fluorinated sites are shown in the upper left corner. The spectra were recorded at 303 K without proton decoupling and referenced relative to TFA. An exponential line broadening of 3 Hz was used to increase signal-to-noise ratio. The arrows in the spectrum of 5FW ECFP indicate two resonances, which were assigned by comparison of peak integrals (see text).

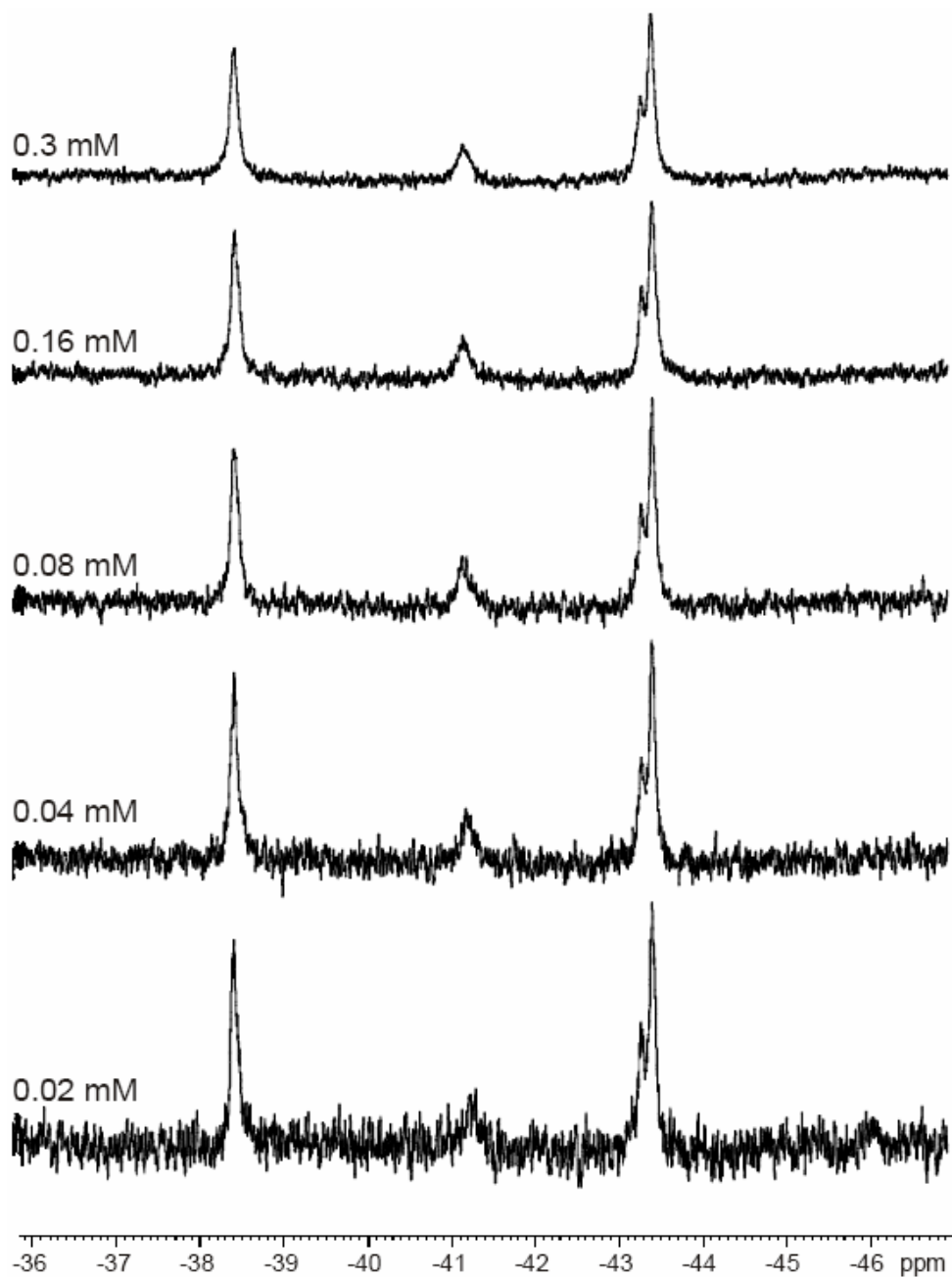


Figure 3.26. Concentration dependence in ^{19}F NMR spectra of 6FW ECFP: protein concentration does not affect the exchanging states in ECFP significantly. This excludes dimerization as source of the observed states. The concentration is given as the total concentration of monomeric protein (27 kDa). The decrease in signal-to-noise ratio results from the linear increase of signal intensity with the number of scans per spectrum but a quadratic increase of noise level.

3.2.2.2 Discussion

The ^{19}F -NMR studies have taken advantage of high sensitivity of ^{19}F chemical shifts to the changes in local environments including van der Waals packing interactions and local electrostatic fields with no background signals from solvent. In most cases, crystallographic studies confirmed that structures of fluorinated proteins are indistinguishable from those of wild-type protein forms (Minks et al. 2000). This makes ^{19}F NMR a powerful technique aromatic residues are especially suitable targets for fluorination because they are normally less likely to be involved in internal motions than smaller amino acids, thus facilitating chemical shift calculations. In this context, the residue specific labeling of EGFP and ECFP with fluorine containing tryptophan analogues offers a good model for NMR studies as well as the opportunity to study possible effects of fluorination on the dynamic properties of protein chromophore. Spectroscopic investigations of substituted proteins at neutral pH clearly indicate that incorporation of electron-withdrawing fluorine in the environment (i.e. W57) of the EGFP chromophore does not affect absorbance and steady-state fluorescence properties of this molecule (Figure 3.24). On the other hand, additional fluorination of W66 of ECFP provides direct integration of the fluorine nucleus into the chromophore. Therefore, the limited blue shift in both fluorescence and absorbance as well as the dramatic decrease in fluorescence intensity upon fluorination can be attributed to this H→F substitution. However, the general behavior of ECFP is not substantially changed by fluorination since two states of chromophore seen as absorption bands at around 430 nm and 450 nm as well as in fluorescence are preserved (Figure 3.24). Thus these spectral properties represent relatively static properties of the GFP chromophore whose dynamic behavior can certainly be better dissected using ^{19}F NMR. Since light absorption takes place on a femtosecond timescale the information on structural dynamics of proteins obtained by UV absorption spectroscopy is limited to an instantaneous average of the immediate environment of the absorbing moiety on this timescale (Laskowski 1995). Compared to the UV spectroscopy NMR is able to characterize motions on a much broader range of timescales. Each of the ECFP residues W57 and W66 exhibits two ^{19}F NMR resonances, which is a clear sign for slow exchange processes. For both tryptophans, W66 and W57, the very low $f = \text{NOE} - 1$ values exclude the existence of pronounced motions on the picosecond to nanosecond timescale. Excessive motions on timescales shorter than picoseconds are also excluded by the small values of T_1 . ^{19}F linewidths indicate the existence of motions on a time scale of nanoseconds and longer. The temperature dependence and denaturation behavior observed

in the ^{19}F NMR spectra of 6FW ECFP provides direct evidence for the existence of slow exchange processes. The mean lifetime of the W66 states was found to be 1.2 ms to 1.4 ms for temperatures in the range of 293 K to 318 K. Slow conformational exchange between two ground states of the chromophore is also in-line with the existence of two UV absorption and fluorescence peaks for CFP and their temperature dependence. The pH titration (Figure 3.13) suggests that the two exchanging states of the chromophore in ECFP cannot be attributed to different protonation states of the indole nitrogen. This leads to the conclusion that the two states of the CFP chromophore are formed in a different way compared to GFP where the protonation state of the phenolic oxygen determines the state of the chromophore. Both tryptophans in ECFP show double peaks. Therefore, it could be thought first that this double peak behavior may be general for tryptophans, but the ^{19}F NMR spectrum of GFPuv showed only a single peak for W57. The more probable explanation of double signals for both tryptophans in the NMR spectrum of ECFP is that the signals of W57 originate from different processes than those of W66, as the resonances of W57 can be easily explained by the exchange between cis and trans proline conformation (Wüthrich 1994) in the P56-W57-P58 sequence. In order to determine the exchange process the steric interactions in the chromophore cavity of CFP has to be examined for which a crystal structure would be desirable. Since attempts to crystallize fluorinated CFP were not yet successful the structure CFP was modeled based on X-ray data from the closely related 4- amino-tryptophan CFP (Budisa et al. 2002). It is well known that even minor substitutions in the core of proteins can cause substantial structural rearrangements therefore the model may give only a first impression on the steric interactions going on in CFP; nevertheless, this provides substantial aid for interpreting the NMR data. In principle, three possibilities for the exchange process within CFP exist that may explain our experimental data: (a) the two states correspond to the monomer-dimer states of CFP, (b) they are related to the exchanging ground states of the chromophore itself or (c) sidechains in the vicinity of the chromophore exchanges between two conformations. The explanation that the two states are simply the monomer and dimer of the protein is the least plausible interpretation. Although GFPs are known to dimerize, the dissociation constant of wild-type GFP is approximately $K_D = 0.1$ mM as measured by analytical ultracentrifugation (Philips 1998). The K_D indicates weak binding for the dimer formation; in addition the symmetric dimers seen in the X-ray structures are rather loose without any interleaved monomer units (Yang et al. 1996) so that even if GFP formed a tight complex the immediate atomic environment,

which to the first approximation is responsible for the difference in the chemical shift, would be almost identical for a given nucleus in the monomer and dimer unit. This is in strong contrast to the large difference in ^{19}F chemical shift ($\Delta\delta = 2.6$ ppm) between the two states, W66A and W66B. Even in the case of a tight nM range dimer (Schatz et al. 2001), a single set of NMR peaks was observed in the ^1H - ^{15}N correlation spectrum, indicating the monomer-dimer equilibrium to be in fast exchange with respect to the NMR time scale. Aromatic ring flipping is a well-known phenomenon found in proteins and to evaluate the possibility of the CFP chromophore being directly involved in such an exchange process, an examination of the known GFP structures is useful. In GFP, the chromophore is stabilized by a hydrogen bond network, which mainly involves its phenolic hydroxyl group and H148 and T203. Despite tight packing in the chromophore cavity it was shown by molecular dynamic simulations that concerted rotations around the τ and ϕ angles, so-called hula-twists, are a possible pathway of photoisomerization of the GFP chromophore since electronic excitation reduces the energy barrier for these rotations (Weber et al. 1999; Chen et al. 2001). Without electronic excitation the hula-twist will occur only with a reduced rate. The model of CFP structure predicts an overall structure similar to the well-known GFP structures, as expected. In CFP, a rotation of the chromophore around the ϕ axis is very unlikely since this would cause a rotation of the long axis of the indole, which is hindered by surrounding residues. But a rotation around the τ axis in concert with an in-plane rotation of the imidazolinone ring appears to be possible (Figure 3.25.), since both rotamers $\tau=0^\circ$ and $\tau=180^\circ$ fit into the same chromophore cavity. Weber et al. showed that the cis and trans isomers of the same protonation state in GFP are similar in their heats of formation (< 5 kJ/mol) (Weber et al. 1999; Chen et al. 2001). Applied to ECFP, this is in agreement with the small energy difference determined by NMR. Weber et al. (1999) also calculated the energy barriers for bond rotations in the ground state of the GFP chromophore. Most of them are in the range of 100 - 200 kJ/mol, except e.g. the ϕ rotation in the neutral and anionic state or the τ rotation in a zwitterionic state, which have lower energy barriers. Since the free energy of activation for the exchange process was estimated to be in the range of 57-63 kJ/mol, the rotation around one of the bonds connecting the two rings of the chromophore appears to be feasible for the indoleimidazolinone chromophore of CFP.

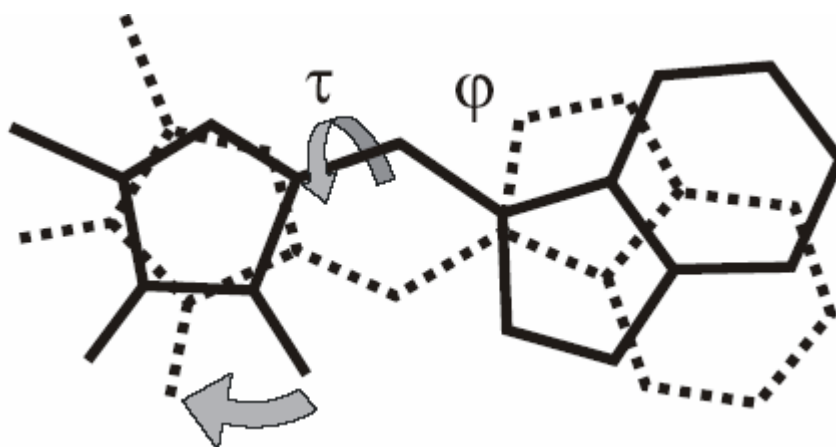


Figure 3.27. The possible rotamer conformations of the CFP chromophore are illustrated. A rotation around τ is accompanied by an in-plane rotation of the imidazolinone ring.

Another possible reason for slow motion processes is the exchange between different conformations of neighboring sidechains. In GFP, two residues, H148 and T203, are closest to the phenolic ring of the chromophore. In most of the crystal structures of GFPs, the β -barrel is somewhat perturbed around H148 with residues 144 to 150 not being hydrogen bonded to the adjacent backbone residues 165 to 170 (Wachter et al. 1998). This suggests that H148 has an increased flexibility compared to other residues within the β -barrel. The comparison of the CFP model structure with the GFP structure suggests that it is sterically feasible for H148 in CFP to adopt two conformations: one with the H148 sidechain oriented towards the chromophore like in GFP and another one with H148 pointing towards the solvent (Figure 3.28).

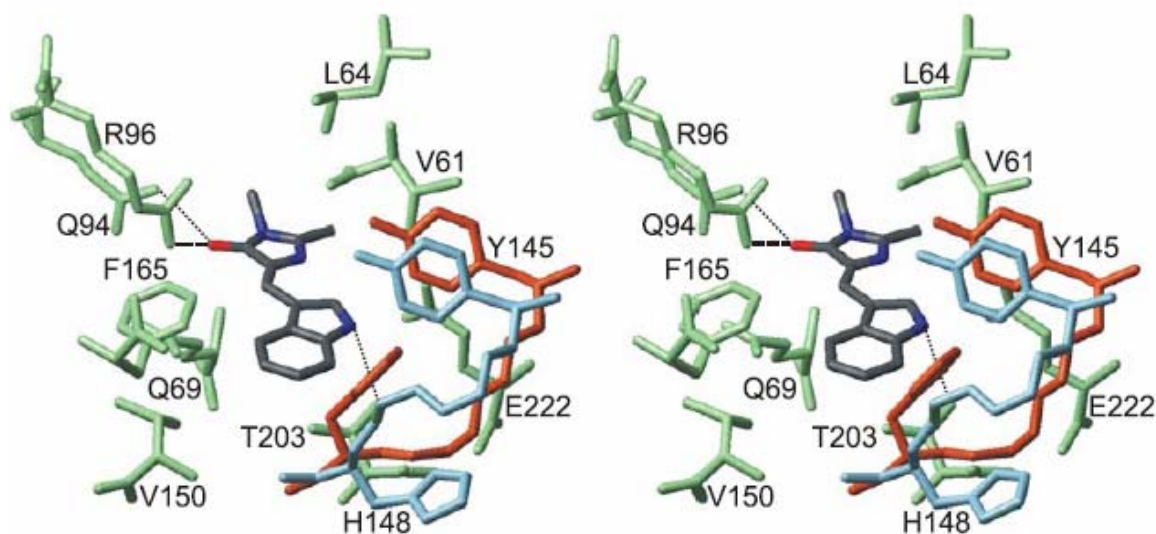


Figure 3.28. Stereo view of the chromophore environment in the CFP model structure: Residues surrounding the chromophore are shown in green. The chromophore itself is shown as a stick representation; the carbon, nitrogen and oxygen atoms are depicted in black, blue and red, respectively. The comparison of the conformation around residue H148 for GFP (in this case taken from PDB entry lemm (Palm et al., 1997)) and the CFP model is shown in orange and cyan, respectively. Close inspection of both structures suggests that the conformation present in GFP may also be adopted in CFP. Hydrogen bonds are depicted as dotted lines.

The different conformations are expected to introduce distortions in the backbone conformation of the adjoining residues. The hydrogen bond between H148 and the chromophore is removed in CFP, at least in one conformation. If the histidine sidechain flips from a chromophore-neighboring conformation to a solvent-exposed conformation, the chemical shift of the fluorine atoms at the 4-, 5- or 6-position of the indole will be affected by the change in electronic environment. A flipped H148 sidechain opens a channel to the solvent and additionally gives the chromophore a larger degree of motional freedom. But the NMR pH-titration clearly shows that the chromophore is not significantly affected by pH changes. Either surrounding hydrophobic residues hinder the entry of water molecules or the imino nitrogen is involved in a hydrogen bond and therefore does not titrate in the measured range of pH. Assuming that the solvent exposed conformation of H148 is higher in energy, the increased disorder associated with H148 flipping from its position in the protein core to a more flexible, solvent-exposed position may account for the observed enthalpy-entropy compensation.

Taken together the exchange of H148 between two conformational states is the most probable explanation for our NMR observations, but more X-ray studies are necessary to provide direct evidence.

3.2.3 *α -PEC (α -subunit of phycoerythrocyanin)*

3.2.3.1 Results

Expression of α -PEC in *E. coli* introduces an effective means of labeling of the apoprotein with ^{13}C , ^{15}N and ^2H isotopes. Figure 3.30 shows the ^{15}N -HSQC NMR spectrum of ^{15}N -labeled α -PEC cloned and expressed in *E. coli*. This allowed us to study the conformations of the chromophore itself and the apoprotein part independently by isotope edited NMR spectroscopy. A long term aim of this research is to study the influence of the chromophore (phycoviolobin – PVB (purified from *Spirulina platensis*)) on the dynamics of the protein structure.



Figure 3.29. SDS-PAGE of fractions from purification (gel filtration Sephadex S-75) of α -PEC after reconstituting with the chromophore. Sample 9 contains a complex of α -PEC and the lyases PecE and PecF. Sample 23 contains α -PEC and α -PEC with a deletion of 14 amino acids from the His-Tag.

The α -PEC production consist of three stages: (a) protein expression and purification, (b) chromophore chemical synthesis and (c) coupling with help of enzymatic proteins (lyases). The initial problem with α -PEC purification was that the classical purification method involving native gel electrophoresis introduces low molecular mass

impurities like, for example, partially polymerized acrylamide. Although undetectable by mass spectroscopy these impurities gave rise to pronounced signals in NMR spectra. We overcame this problem using gel filtration (S75 – Sephadex) instead of preparative electrophoresis and we were able to separate α -PEC coupled with chromophore from lyases. The resulting samples were 95 – 98% pure as proven by mass spectroscopy and initial crystallization attempts.

This highly purified sample was exhibiting autodigestion activity (Figure 3.29) as also reported by others (Reuter et al. 1993). Mass spectroscopy indicated that the first 14 amino acids from the N – terminus (His – tag, plus protein sequence) were removed (first N-terminus α -helix). Two first helices of the α -PEC are known to be responsible for both protein oligomerization in the photosystem (Reuter et al. 1993) and α -PEC protein aggregation in solution (Reuter et al. 1993; Storf et al. 2001; Wiegand et al. 2002). Our results show that the presence of first of two N-terminus helices inhibits oligomerization. When the second helix becomes exposed to solution it mediates aggregation. The initial samples were stable and were of “NMR quality” for approximately three weeks. Without the 14 N-terminal amino acids the protein forms aggregates. This aggregation leads to broadening of NMR lines and in some cases to precipitation of the protein. Further optimization of the α -PEC construct is required for production of long-term stable, NMR samples. It has been shown that N-terminal degradation does not influence neither I nor II photochemistry which remains identical, qualitatively and quantitatively, to that of a native α -PEC (Zhao et al. 2003). On the other hand the C-terminally shortened protein could not be reconstituted with lyases, probably due to misfolding (Zhao et al. 2003).

As stated in the introduction α -PEC has two types of photochemistry: by irradiating with light of wave length 500 or 577 nm it could be switched between Z - α I and E - α I respectively (first photochemistry) and by reduction or oxidation of Cys 98 and 99 between α I and α II (second photochemistry) (Hong et al. 1993). Irradiation was performed as reported by Hong et al. (1993). To investigate dynamics of the first photoconversion we recorded a series of ^{15}N – HSQCs (Figure 3.30 and Table 3.2).

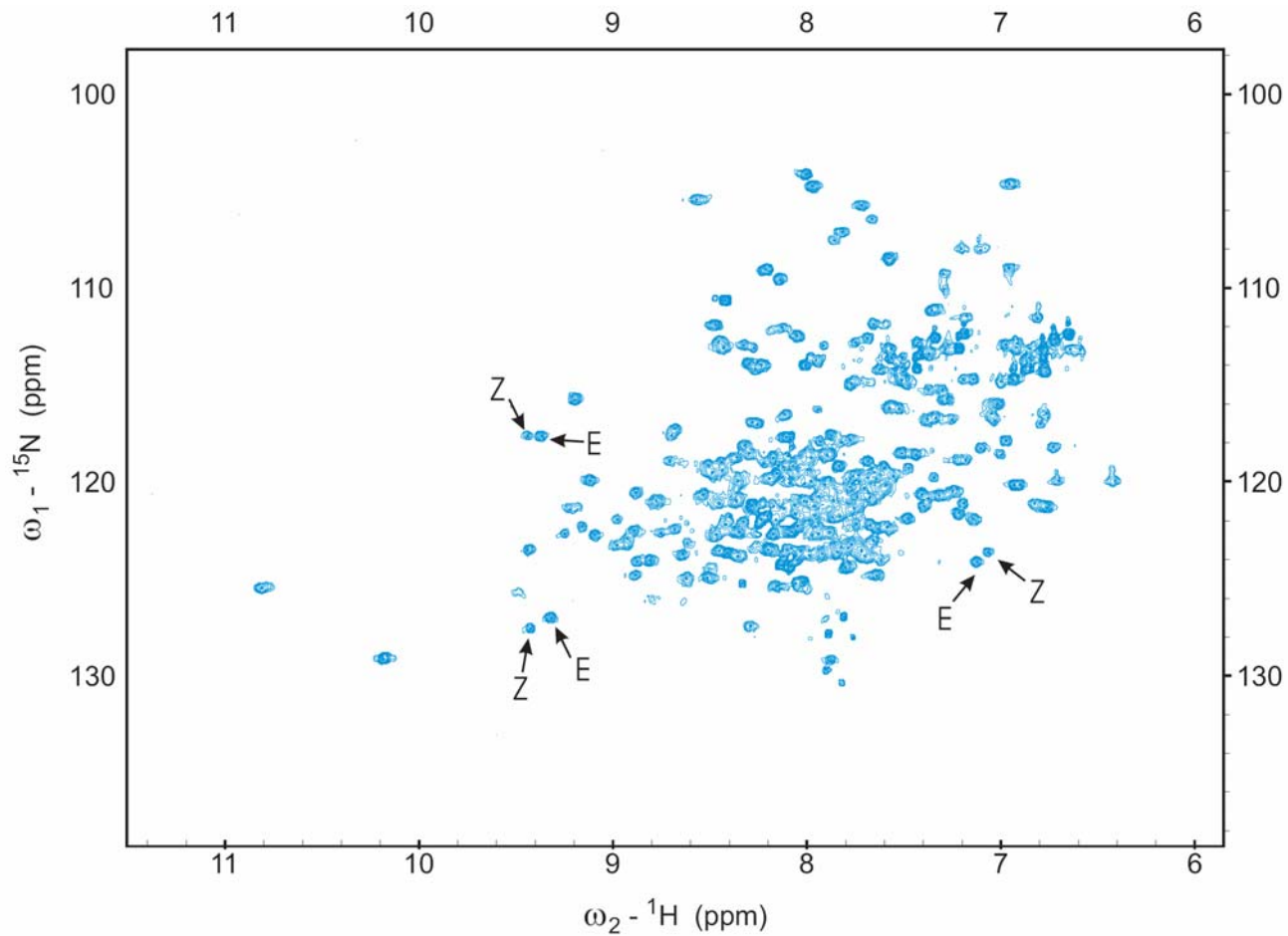


Figure 3.30. The ^{15}N -HSQC NMR spectrum of ^{15}N -labeled α -PEC coupled with the chromophore in phosphate buffer pH 7.00 and at 300 K. Arrows indicate double peaks of E and Z states of the holoprotein.

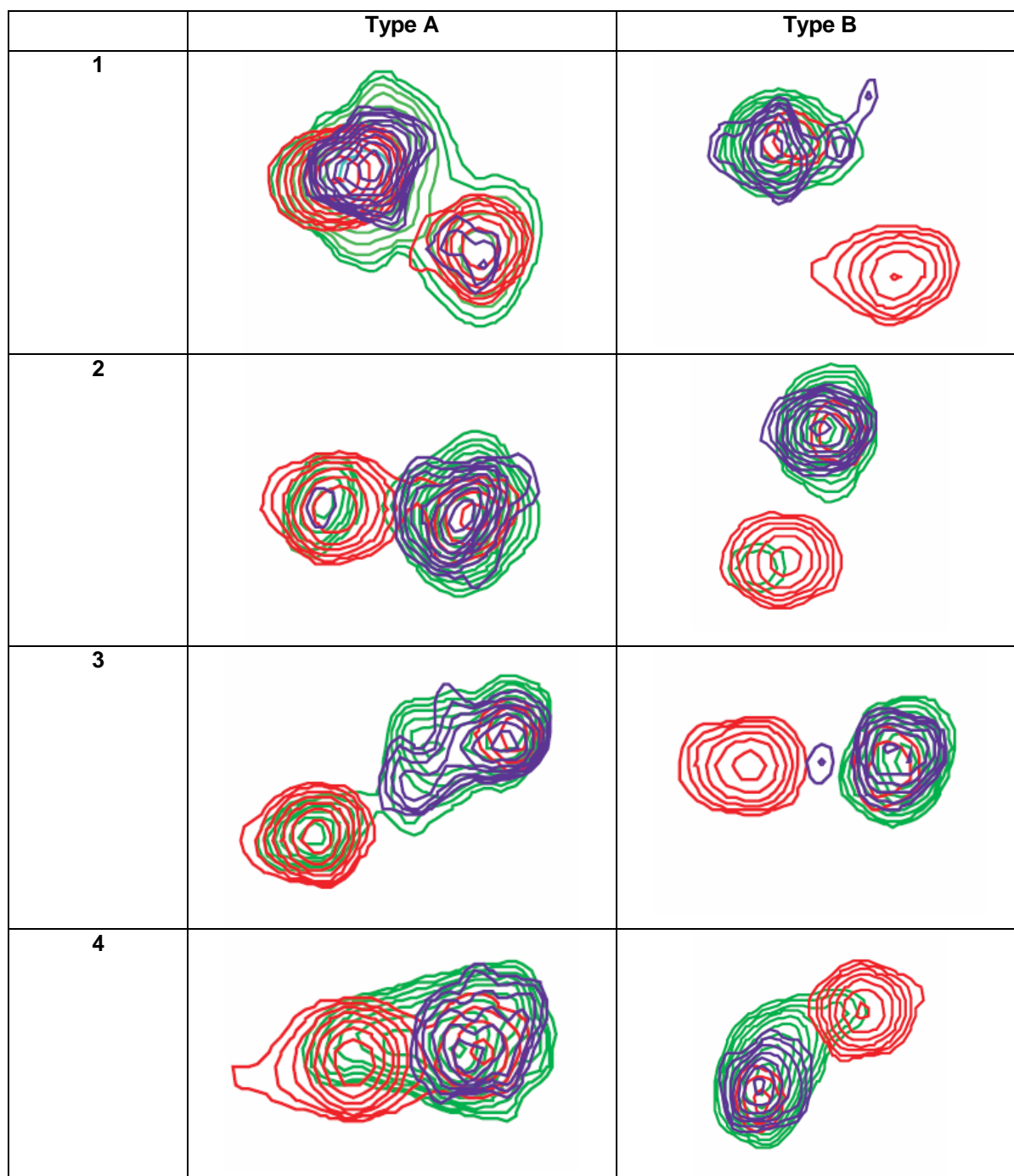


Table 3.2. A close-up of some of the peaks from ^{15}N – HSQC spectrum of ^{15}N – labeled α -PEC holoprotein affected by $Z \rightarrow E$ photoconversion. Each peak represents one amino acid residue. Peaks were divided into two groups type A and B. In type A peaks were mostly doublets after overnight irradiation with 500 nm light (so called “Z state”) (red), after switching to the E state (irradiation with 577 nm light) (violet) they formed singlets. After 5 hours in the darkness they partially relaxed back to the previous position (green). In type B after excitation to E state recovery to low energy “Z state” is much slower and is practically absent even after 5 hours in the darkness.

The ^{15}N – HSQC spectrum of the ^{15}N – labeled α -PEC has peaks that were affected by the E – Z photoconversion (Figure 3.30). We observed that the ^{15}N – HSQC of α -PEC after irradiation with 500 nm light (“Z form”) exhibits double peaks. This does not change with time (5 hours in darkness). The E form spectrum displays mostly single peaks. Spectra of E form taken after 5 hours show that E form tends to interconvert to Z form with time, even though the sample is kept in darkness (Table 3.2). Peaks of the E form could be classified into two classes due to the rate of interconversion from E to Z (Table 3.2). The type A is characterized by a fast E \rightarrow Z interconversion process which in ca. 5 hours leads to the formation of double peaks. The type B displays slow interconversion and in 5 hours nearly no double peaks appear. Measurements taken parallelly on the UV/VIS spectrometer showed no difference between freshly irradiated E form and the E form after irradiation kept for 5 hours in the darkness. Efficiency of photoconversion was ensured by applying 500 nm (Z form) or 577 nm (E form) light overnight. Both samples were checked by UV/VIS spectrometry and were identified as a fully converted to respective forms (Figure 3.31).

Structural analysis of de novo, in-vitro synthesized phytochromobilins has also been carried out (Figure 3.32). Novel chromophores would form new photodynamic systems upon binding to α -PEC. The structural investigation of newly designed and synthesized chromophores is therefore required for further analysis of the protein-phytochromobilin photosystem.

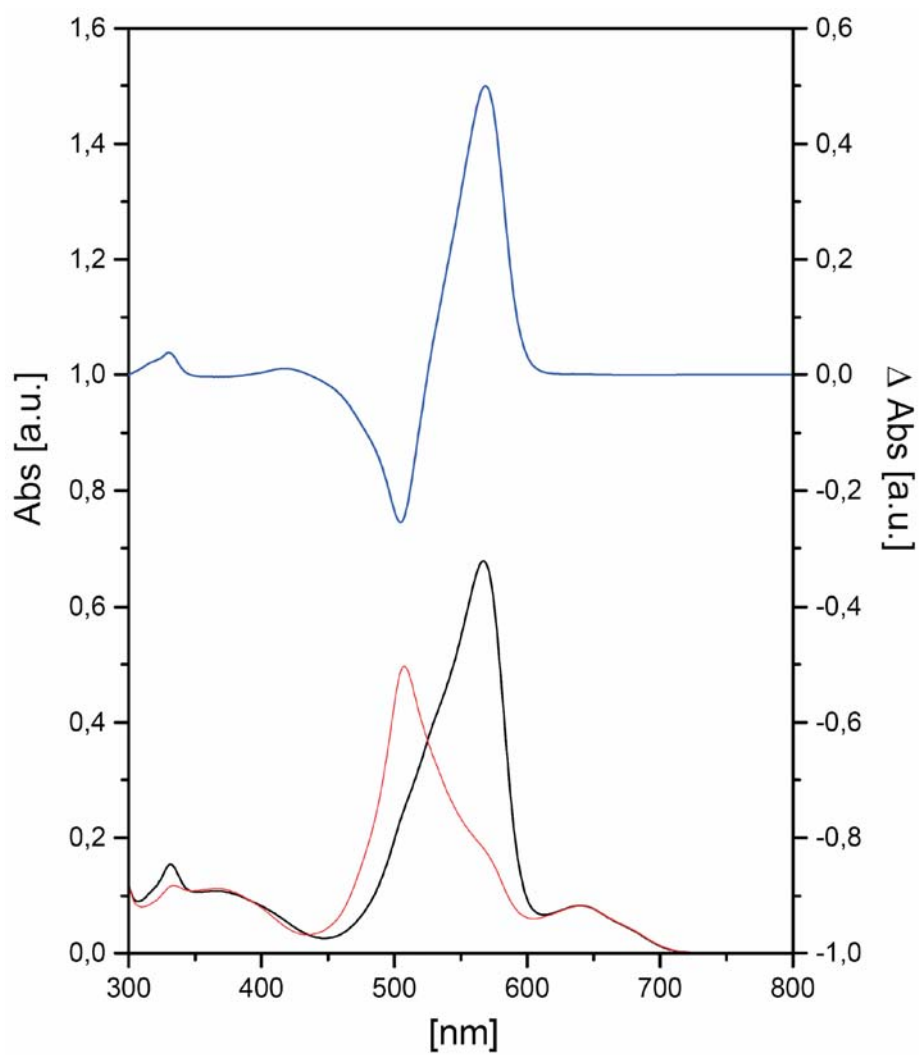


Figure 3.31. UV/VIS absorbance spectrum of E (red) and Z (black) form of α -PEC. Blue line shows difference spectrum.

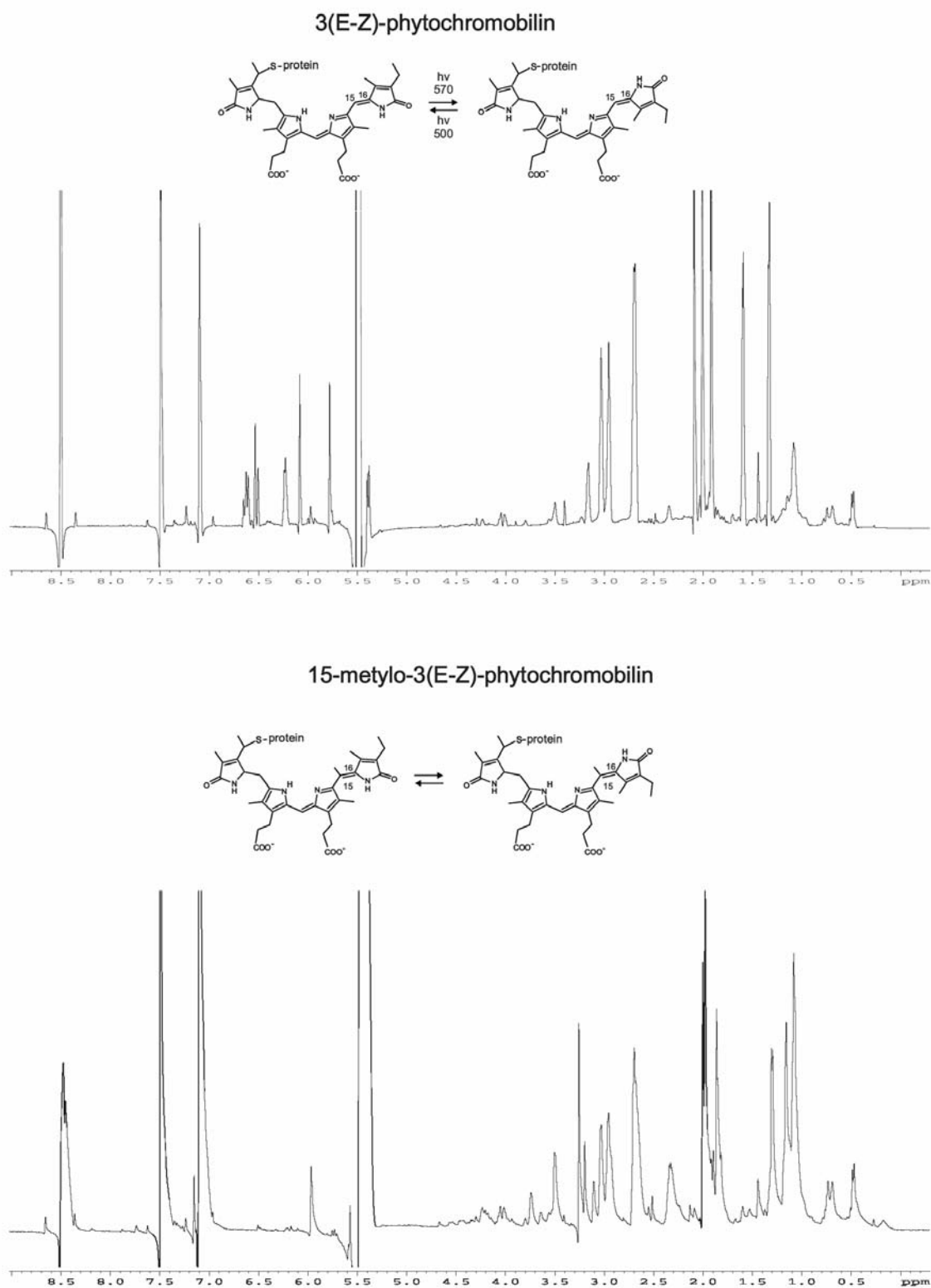


Figure 3.32. 1D-proton spectrum of 3(E-Z)-phytochromobilin and 15-metylo-3(E-Z)-phytochromobilin; chromophores of α -PEC photosystem. Spectra recorded on highly concentrated samples in pyridine d^5 on a 600 MHz NMR spectrometer at 298 K.

3.2.3.2 Discussion

A widely accepted model says that E – Z phototransformation of the chromophore involves isomerization around the $\Delta 15, 16$ double bond between the pyrrole rings C and D (Bishop et al. 1987; Neff et al. 2000). The influence of different configurations of the chromophore on the conformational behavior of the protein moiety was the subject of our investigation.

Double peaks observed in the ^{15}N -HSQC indicate existence of two distinct conformations of the protein in what is usually called “form Z of holoprotein”. Protein in this “Z state” which is characterized by maximum of absorbance at 566 nm (Figure 3.31) is actually a mixture of holoprotein with chromophore in the Z (approximately 70%) and E (30%) (Table 3.2). Existence of double peaks could not be explained by dimerization or oligomerization (Schatz et al. 2001). Experimental data suggest also conformational instability of E form. ^{15}N -HSQC spectra acquired directly after irradiation and acquired after irradiation followed by 5 hours in darkness differ. Parallel measurements taken on the UV/VIS spectrometer show no difference between these two samples. It has been reported that both E and Z state are stable in darkness (Hong et al. 1993).

It is known that the chromophore of α -PEC in the E form is more flexible (Wiegand et al. 2002) and the protein chain is more rigid than in the Z form. It could be that both E and Z forms have some degree of unhomogeneity in respect to the conformation on the protein. In addition to this E form is undergoing dynamic changes of protein conformation due to interconversion of some of the residues.

Further research on dynamics of both forms of α -PEC is required to understand the chromophore photoconversion mechanism and the role of protein in this process.

4 SUMMARY

This thesis is based on the studies that have been carried out in the Department of Structural Research at the Max Planck Institute for Biochemistry from December 2000 to January 2004. Two projects were undertaken. The first was the study of function and interactions of cell cycle proteins and cell differentiation proteins. The proteins studied were retinoblastoma protein (pRb), cyclin-dependent kinase 6 (CDK6) and cell differentiation factors MyoD and Id-2. The second project focused on structural and dynamic investigations on photoactive proteins green fluorescent protein (GFP), cyan fluorescent protein (CFP) and alpha subunit of phycoerythrocyanin (α -PEC).

Cell growth and differentiation require precise interplay between cell cycle and differentiation factors. Because much of the data about the interactions between proteins crucial for this interplay is contradictory, the aim of the study was to find whether pivotal cell cycle proteins (pRb, and CDK6) can interact directly with myogenic differentiation promoting (MyoD) and inhibiting (Id-2) factors. These potential interactions were tested with pull down assays, gel filtration, mass spectroscopy, and the nuclear magnetic resonance (NMR) spectroscopy. We found that there are no direct protein-protein interactions *in vitro*. Indirect modulation of gene expression level of these proteins is therefore proposed. In addition we have also shown that the C-terminal part of human and chicken MyoD does not interact with human cycline-dependent kinase 6, contrary to the recent report about such a interaction taking place.

The NMR characterization of the GFP and CFP created an opportunity for investigation of their dynamic properties. Newly adopted method of *in vitro* protein production allowed us to make ^{15}N -His labeled samples of GFP without any cross-labeling to other amino acids. This probe allowed for identification of peaks corresponding to the important residues that interact with the chromophore (e.g. His148). The ^{15}N - HSQC NMR spectra of the H148G GFP mutant and selectively labeled histidine samples indicated the presence of two conformations of the protein in slow exchange on the NMR time scale. Also the ^{19}F NMR studies of variants of GFP and CFP labeled with fluorinated tryptophans, supported by temperature-, concentration- and folding-dependent experiments, provided direct evidence for the existence of a slow exchange process between two different conformational states of CFP.

Photoisomerization of holoprotein α -PEC involves E – Z isomerization around a double bond in the chromophore. In our research we could observe how the two states of the chromophore influence the structure of protein. Using ^{15}N – HSQC we observed double peaks that indicate existence of both E and Z conformations in α -PEC in the 566 nm form (made by 500 nm light irradiation). Moreover our experimental data suggest that the E form (induced by 577 nm irradiation) is conformationally unstable. Both the E and Z forms have some degree of inhomogeneity with respect to the conformation of the protein. We found additionally that the E form undergoes dynamic changes in protein conformation due to interconversion to the Z form, even in darkness. This process does not affect UV/VIS spectroscopic properties of form E but it certainly affects the structure of the protein.

5 ZUSAMMENFASSUNG

Die Arbeit basiert auf Untersuchungen, welche in der NMR-Gruppe der Abteilung für Strukturuntersuchungen am Max-Planck-Institut für Biochemie vom Dezember 2000 bis Januar 2004 durchgeführt wurden. Zwei Projekte wurden bearbeitet: (a) die Untersuchung von Funktions- und Interaktionsmustern von bedeutenden Zellzyklus- (retinoblastoma protein (pRb) und cyclin-dependent kinase 6 (CDK6)) und Zelldifferenzierungs- (MyoD und Id-2) proteinen; (b) strukturelle und dynamische Untersuchungen an den photoaktiven Proteinen GFP (green fluorescent protein), CFP (cyan fluorescent protein) und α -PEC (alpha-Untereinheit von Phycoerythrocyanin).

Zellwachstum und Differenzierung erfordern ein präzises Zusammenspiel zwischen Zellzyklus- und Differenzierungsfaktoren. Da es eine Menge sich widersprechender Daten zur Wechselwirkung zwischen den für dieses Zusammenspiel entscheidenden Proteinen gibt, war das Ziel der Untersuchung, herauszufinden, ob die Zellzyklusproteine (pRb und CDK6) eine direkte Protein-Protein-Wechselwirkungen mit den myogenischen differenzierungsfördernden (MyoD) bzw. -hemmenden (Id-2) Faktoren haben. Wir untersuchten diese Möglichkeiten unter Verwendung von „Pull-Down“-Experimenten, Gel-Filtration, Massenspektroskopie, und kernmagnetische Resonanzspektroskopie (NMR). Wir fanden heraus, dass es *in vitro* keine direkten Protein-Protein-Wechselwirkungen gibt. Weiterhin haben wir gezeigt, dass die C-terminale Domäne von humanem und Hühner-MyoD nicht mit humanem CDK6 wechselwirkt.

Die Charakterisierung von GFP und CFP mit NMR eröffnete die Möglichkeit für Untersuchungen ihrer dynamischen Eigenschaften. Die neu eingeführte Methode der *in vitro*-Proteinproduktion erlaubte uns, ^{15}N -Histidin markierte Proben von GFP herzustellen, die frei von Kreuzmarkierungen durch andere Aminosäuren waren wie sie bei der *in vivo* Markierung vorkommen. Mit diesen Proben konnten die Peaks ermittelt werden, die für die Wechselwirkung mit dem Chromophor wichtig sind (z.B. His148). Die ^{15}N -HSQC NMR-Spektren der H148G GFP-Mutant, die selektiv ^{15}N -Histidin markiert war, deuten auf die Präsenz von zwei Konformationen des Proteins, in einem langsamen gegenseitigen Austausch auf der NMR-Zeitskala, hin. Die ^{19}F -NMR-Studien mit Varianten von GFP und CFP, die mit fluorierten Tryptophanen markiert wurden lieferten, unterstützt durch temperatur-, konzentrations- und faltungsabhängige Experimente, direkte Hinweise auf die

Existenz eines langsamen Austauschprozesses zwischen zwei verschiedenen konformativen Zuständen des CFP.

Die Photoisomerisierung des Holoproteins α -PEC beinhaltet eine E-Z-Isomerisierung an einer Doppelbindung im Chromophor. Während unserer Untersuchungen konnten wir beobachten, wie die verschiedenen Zustände des Chromophors die Struktur des Proteins beeinflussen. In ^{15}N -HSQC Experimenten konnten wir Doppel-Peaks beobachten, welche die Existenz beider Konformationen E und Z in α -PEC in der 566 nm Form zeigt. (Die 566 nm Form entsteht durch Belichten des Proteins mit einer Wellenlänge von 500 nm) Weiterhin weisen unsere experimentellen Daten auf eine konformative Instabilität der E-Form hin die bei Bestrahlung mit einer Wellenlänge von 577 nm entsteht.

Es könnte sein, dass sowohl die E- als auch die Z-Form einen gewissen Grad an Inhomogenität bzgl. der Konformation des Proteins besitzen. Darüber geht die E-Form in einer dynamischen Veränderung aufgrund von Interkonversion in die Z-Form über. Dieser Prozess beeinflusst nicht die UV/VIS-spektroskopischen Eigenschaften der E-Form, aber er beeinflusst mit Sicherheit die Struktur des Proteins.

6 REFERENCES

- Alesse, E., F. Zazzeroni, A. Angelucci, G. Giannini, L. Di Marcotullio and A. Gulino (1998). "The growth arrest and downregulation of c-myc transcription induced by ceramide are related events dependent on p21 induction, Rb underphosphorylation and E2F sequestering." *Cell Death Differ* 5: 381-9.
- Arnold, H. H. and T. Braun (1996). "Targeted inactivation of myogenic factor genes reveals their role during mouse myogenesis: A review." *Int J Dev Biol* 40: 345-353.
- Arnold, H. H. and B. Winter (1998). "Muscle differentiation: more complexity to the network of myogenic regulators." *Curr Opin Genet Dev* 8: 539-544.
- Beatz, M. (1997). "One century of protein crystallography - the phycobiliprotein." *Biol. Chem.* 378: 167-176.
- Benezra, R., R. L. Davis, D. Lockshon, D. L. Turner and H. Weintraub (1990). "The Protein Id - a Negative Regulator of Helix-Loop-Helix DNA-Binding Proteins." *Cell* 61: 49-59.
- Betton, J. M. (2003). "Rapid translation system (RTS): a promising alternative for recombinant protein production." *Curr Protein Pept Sci* 4: 73-80.
- Biggs, J., E. V. Murphy and M. A. Israel (1992). "A Human Id-Like Helix Loop Helix Protein Expressed During Early Development." *Proc Natl Acad Sci U S A* 89: 1512-1516.
- Bishop, J. (1996). *Molecular Oncology*. New York, Scientific American Inc.
- Bishop, J. E., H. Rapoport, A. V. Klotz, C. F. Chan, A. N. Glazer, P. Fuglistaller and H. Zuber (1987). "Chromopeptides from Phycoerythrocyanin - Structure and Linkage of the 3 Bilin Groups." *J Am Chem Soc* 109: 875-881.
- Björn, L. (1979). "Photoreversibly photochromic pigments in organism: properties and role in physiological light perception." *Quart Rev Biophysic* 12: 95-113.
- Blattner, C., A. Sparks and D. Lane (1999). "Transcription factor E2F-1 is upregulated in response to DNA damage in a manner analogous to that of p53." *Mol Cell Biol* 19: 3704-13.
- Bryant, D. (1982). "Phycoerythrocyanin and phycoerythrin: Properties and occurrence in cyanobacteria." *J Gen Microbiol* 128: 835-844.
- Budisa, N., M. Rubini, J. H. Bae, E. Weyher, W. Wenger, R. Golbik, R. Huber and L. Moroder (2002). "Global replacement of tryptophan with aminotryptophans

- generates non-invasive protein-based optical pH sensors." *Angew Chem Int Ed Engl* 41: 4066-9.
- Campbell, R. E., O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, D. A. Zacharias and R. Y. Tsien (2002). "A monomeric red fluorescent protein." *Proc Natl Acad Sci U S A* 99: 7877-82.
- Canet, D., A. M. Last, P. Tito, M. Sunde, A. Spencer, D. B. Archer, C. Redfield, C. V. Robinson and C. M. Dobson (2002). "Local cooperativity in the unfolding of an amyloidogenic variant of human lysozyme." *Nat Struct Biol* 9: 308-15.
- Cantor, C., Schimmel, PR. (1980). *Biophysical chemistry part II: Techniques for study of biological structure and function*. San Francisco, W.H. Freeman and Company.
- Cavenee, W. K. and R. L. White (1995). "The Genetic-Basis of Cancer." *Sci Am* 272: 72-79.
- Cenciarelli, C., F. De Santa, P. L. Puri, E. Mattei, L. Ricci, F. Bucci, A. Felsani and M. Caruso (1999). "Critical role played by cyclin D3 in the MyoD-mediated arrest of cell cycle during myoblast differentiation." *Mol Cell Biol* 19: 5203-17.
- Chen, M. C., C. R. Lambert, J. D. Urgitis and M. Zimmer (2001). "Photoisomerization of green fluorescent protein and the dimensions of the chromophore cavity." *Chem Phys* 270: 157-164.
- Choi, J., Costa, ML., Mermelstein, CS., Chagas, C., Holtzer, S., Holtzer, H. (1990). "MyoD converts primary dermal fibroblasts, chondroblasts, smooth muscle, and retinal pigmented epithelial cells into striated mononucleated myoblasts and multinucleated myotubes." *Proc Natl Acad Sci U S A* 87: 7988-92.
- Christy, B. A., L. K. Sanders, L. F. Lau, N. G. Copeland, N. A. Jenkins and D. Nathans (1991). "An Id-Related Helix Loop Helix Protein Encoded by a Growth Factor-Inducible Gene." *Proc Natl Acad Sci U S A* 88: 1815-1819.
- Cieslar, C., T. A. Holak and H. Oschkinat (1990). "A Program for the Evaluation of 3d Spectra Applied to the Sequential Assignment of Bpti Utilizing 3d Tocsy-Noesy." *J Magn Reson* 87: 400-407.
- Crescenzi, M., T. P. Fleming, A. B. Lassar, H. Weintraub and S. A. Aaronson (1990). "MyoD induces growth arrest independent of differentiation in normal and transformed cells." *Proc Natl Acad Sci U S A* 87: 8442-6.
- Czisch, M. and R. Boelens (1998). "Sensitivity enhancement in the TROSY experiment." *J Magn Reson* 134: 158-60.

-
- Davis, R., Weintraub, H., Lassar, AB. (1987). "Expression of a single transfected cDNA converts fibroblasts to myoblasts." *Cell* 51: 987-1000.
- DeCaprio, J. A., J. W. Ludlow, J. Figge, J. Y. Shew, C. M. Huang, W. H. Lee, E. Marsilio, E. Paucha and D. M. Livingston (1988). "SV40 large tumor antigen forms a specific complex with the product of the retinoblastoma susceptibility gene." *Cell* 54: 275-83.
- Dempsey, C. E. (2001). "Hydrogen exchange in peptides and proteins using NMR-spectroscopy." *Prog NMR Spectrosc* 39: 135-170.
- Diercks, T., M. Coles and H. Kessler (2001). "Applications of NMR in drug discovery." *Curr Opin Chem Biol* 5: 285-91.
- Dingley, A. J., J. P. Mackay, B. E. Chapman, M. B. Morris, P. W. Kuchel, B. D. Hambly and G. F. King (1995). "Measuring Protein Self-Association Using Pulsed-Field-Gradient Nmr-Spectroscopy - Application to Myosin Light-Chain-2." *J Biomol NMR* 6: 321-328.
- Dobson, C. M. (2003). "Protein folding and misfolding." *Nature* 426: 884-90.
- Duerring, M., R. Huber, W. Bode, R. Ruembeli and H. Zuber (1990). "Refined three-dimensional structure of phycoerythrocyanin from the cyanobacterium *Mastigocladus laminosus* at 2.7 Å." *J Mol Biol* 211: 633-44.
- Ekhholm, S. V. and S. I. Reed (2000). "Regulation of G(1) cyclin dependent kinases in the mammalian cell cycle." *Curr Opin Cell Biol* 12: 676-684.
- Elledge, S. J. (1996). "Cell cycle checkpoints: Preventing an identity crisis." *Science* 274: 1664-1672.
- Ellmeier, W., A. Aguzzi, E. Kleiner, R. Kurzbauer and A. Weith (1992). "Mutually Exclusive Expression of a Helix Loop Helix Gene and N-Myc in Human Neuroblastomas and in Normal Development." *EMBO J* 11: 2563-2571.
- Enzenauer, C., G. Mengus, A. Lavigne, I. Davidson, H. Pfister and M. May (1998). "Interaction of human papillomavirus 8 regulatory proteins E2, E6 and E7 with components of the TFIID complex." *Intervirology* 41: 80-90.
- Ephrussi, A., G. M. Church, S. Tonegawa and W. Gilbert (1985). "B-Lineage Specific Interactions of an Immunoglobulin Enhancer with Cellular Factors In vivo." *Science* 227: 134-140.

- Ericson, K. K., D. Krull, P. Slomiany and M. J. Gressel (2003). "Expression of cyclin-dependent kinase 6, but not cyclin-dependent kinase 4, alters morphology of cultured mouse astrocytes." *Mol Cancer Res* 1: 654-64.
- Fan, G., X. Ma, B. T. Kren and C. J. Steer (1996). "The retinoblastoma gene product inhibits TGF-beta1 induced apoptosis in primary rat hepatocytes and human HuH-7 hepatoma cells." *Oncogene* 12: 1909-19.
- Fernandez, C. and G. Wider (2003). "TROSY in NMR studies of the structure and function of large biological macromolecules." *Curr Opin Struct Biol* 13: 570-80.
- Fiaux, J., E. B. Bertelsen, A. L. Horwich and K. Wuthrich (2002). "NMR analysis of a 900K GroEL GroES complex." *Nature* 418: 207-11.
- Fielding, L. (2003). "NMR methods for the determination of protein-ligand dissociation constants." *Curr Top Med Chem* 3: 39-53.
- Ford, C. F., I. Suominen and C. E. Glatz (1991). "Fusion tails for the recovery and purification of recombinant proteins." *Protein Expr Purif* 2: 95-107.
- Föster, T. (1950). *Z. Elektrochem* 54: 42-46.
- Fukuda, H., M. Arai and K. Kuwajima (2000). "Folding of green fluorescent protein and the cycle3 mutant." *Biochemistry* 39: 12025-32.
- Georgescu, J., T. Rehm, J. Wiehler, B. Steipe and T. A. Holak (2003). "Backbone H(N), N, C(alpha) and C(beta) assignment of the GFPuv mutant." *J Biomol NMR* 25: 161-2.
- Gietz, R. D., B. Triggs-Raine, A. Robbins, K. C. Graham and R. A. Woods (1997). "Identification of proteins that interact with a protein of interest: applications of the yeast two-hybrid system." *Mol Cell Biochem* 172: 67-79.
- Glazer, A. N. (1985). "Light Harvesting by Phycobilisomes." *Annu Rev Biophys Biophys Chem* 14: 47-77.
- Goddard, T., Kneller, DG. (2001). *Sparky 3*. San Francisco, University of California.
- Gu, W., J. W. Schneider, G. Condorelli, S. Kaushal, V. Mahdavi and B. Nadal-Ginard (1993). "Interaction of myogenic factors and the retinoblastoma protein mediates muscle cell commitment and differentiation." *Cell* 72: 309-24.
- Guan, K. L., C. W. Jenkins, Y. Li, M. A. Nichols, X. Wu, C. L. O'Keefe, A. G. Matera and Y. Xiong (1994). "Growth suppression by p18, a p16INK4/MTS1- and p14INK4B/MTS2-related CDK6 inhibitor, correlates with wild-type pRb function." *Genes Dev* 8: 2939-52.

- Guex, N. and M. C. Peitsch (1997). "SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling." *Electrophoresis* 18: 2714-23.
- Halevy, O., B. G. Novitch, D. B. Spicer, S. X. Skapek, J. Rhee, G. J. Hannon, D. Beach and A. B. Lassar (1995). "Correlation of terminal cell cycle arrest of skeletal muscle with induction of p21 by MyoD." *Science* 267: 1018-21.
- Harbour, J. W. and D. C. Dean (2000). "The Rb/E2F pathway: expanding roles and emerging paradigms." *Genes Dev* 14: 2393-2409.
- Harbour, J. W., R. X. Luo, A. Dei Santi, A. A. Postigo and D. C. Dean (1999). "Cdk phosphorylation triggers sequential intramolecular interactions that progressively block Rb functions as cells move through G1." *Cell* 98: 859-69.
- Haupts, U., S. Maiti, P. Schwille and W. W. Webb (1998). "Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy." *Proc Natl Acad Sci U S A* 95: 13573-8.
- Heichman, K. A. and J. M. Roberts (1994). "Rules to Replicate By." *Cell* 79: 557-562.
- Helms, V., T. P. Straatsma and J. A. McCammon (1999). "Internal dynamics of green fluorescent protein." *J Phys Chem B* 103: 3263-3269.
- Henthorn, P., M. Kiledjian and T. Kadesch (1990). "2 Distinct Transcription Factors That Bind the Immunoglobulin Enhancer Mu-E5/Ke2 Motif." *Science* 247: 467-470.
- Hiebert, S. W., G. Packham, D. K. Strom, R. Haffner, M. Oren, G. Zambetti and J. L. Cleveland (1995). "E2F-1:DP-1 induces p53 and overrides survival factors to trigger apoptosis." *Mol Cell Biol* 15: 6864-74.
- Hino, M., Y. Shinohara, K. Kajimoto, H. Terada and Y. Baba (2002). "Requirement of continuous transcription for the synthesis of sufficient amounts of protein by a cell-free rapid translation system." *Protein Expr Purif* 24: 255-9.
- Hong, Q. and Q. Yang (1993). "Quantum theory of spin waves for magnetic-overlayer systems." *Physical Review B Condensed Matter* 47: 7897-7902.
- Hong, Q., K. H. Zhao and H. Scheer (1993). "2 Different Types of Photochemistry in Phycoerythrocyanin Alpha-Subunit." *Photochem Photobiol* 58: 745-747.
- Hopf, M., W. Gohring, A. Ries, R. Timpl and E. Hohenester (2001). "Crystal structure and mutational analysis of a perlecan-binding fragment of nidogen-1." *Nat Struct Biol* 8: 634-40.
- Iannone, M. (1999). "Exploring Fourier transform techniques with MathCad." *J Chem Educ* 76.

-
- Iavarone, A., P. Garg, A. Lasorella, J. Hsu and M. A. Israel (1994). "The Helix-Loop-Helix Protein Id-2 Enhances Cell-Proliferation and Binds to the Retinoblastoma Protein." *Genes Dev* 8: 1270-1284.
- Jen, Y., K. Manova and R. Benezra (1996). "Expression patterns of Id1, Id2, and Id3 are highly related but distinct from that of Id4 during mouse embryogenesis." *Dev Dyn* 207: 235-252.
- Jen, Y., K. Manova and R. Benezra (1997). "Each member of the Id gene family exhibits a unique expression pattern in mouse gastrulation and neurogenesis." *Dev Dyn* 208: 92-106.
- Jogi, A., P. Persson, A. Grynfeld, S. Pahlman and H. Axelson (2002). "Modulation of basic helix-loop-helix transcription complex formation by Id proteins during neuronal differentiation." *J Biol Chem* 277: 9118-26.
- Juranic, N., V. A. Likic, F. G. Prendergast and S. Macura (1996). "Protein-solvent hydrogen bonding studied by NMR (1)J(NC') coupling constant determination and molecular dynamics simulations." *J Am Chem Soc* 118: 7859-7860.
- Kiledjian, M., L. K. Su and T. Kadesch (1988). "Identification and Characterization of 2 Functional Domains within the Murine Heavy-Chain Enhancer." *Mol Cell Biol* 8: 145-152.
- King, K. L. and J. A. Cidlowski (1998). "Cell cycle regulation and apoptosis." *Annu Rev Physiol* 60: 601-17.
- Klamt, C., E. Knust, K. Tietze and J. A. Camposortega (1989). "Closely Related Transcripts Encoded by the Neurogenic Gene-Complex Enhancer of Split of *Drosophila-Melanogaster*." *EMBO J* 8: 203-210.
- Knudsen, E. S. and J. Y. Wang (1997). "Dual mechanisms for the inhibition of E2F binding to RB by cyclin-dependent kinase-mediated RB phosphorylation." *Mol Cell Biol* 17: 5771-83.
- Kummer, A. D., J. Wiehler, H. Rehaber, C. Kompa, B. Steipe and M. E. Michel-Beyerle (2000). "Effects of threonine 203 replacements on excited-state dynamics and fluorescence properties of the green fluorescent protein (GFP)." *J Phys Chem B* 104: 4791-4798.
- Lakowicz, J. (1999). *Principles of fluorescence spectroscopy*, 2nd ed. New York, Kluwer Academic.

-
- Laskowski, R. (1995). "SURFNET: a program for visualizing molecules surfaces, cavities, and intermolecular interactions." *Mol Graphics* 13: 323-330.
- Lasorella, A., A. Iavarone and M. A. Israel (1996). "Id2 specifically alters regulation of the cell cycle by tumor suppressor proteins." *Mol Cell Biol* 16: 2570-2578.
- Lasorella, A., M. Nosedà, M. Beyna and A. Iavarone (2000). "Id2 is a retinoblastoma protein target and mediates signalling by Myc oncoproteins." *Nature* 407: 592-598.
- Lassar, A. B., J. N. Buskin, D. Lockshon, R. L. Davis, S. Apone, S. D. Hauschka and H. Weintraub (1989). "MyoD Is a Sequence-Specific DNA-Binding Protein Requiring a Region of Myc Homology to Bind to the Muscle Creatine-Kinase Enhancer." *Cell* 58: 823-831.
- Lassar, A. B., R. L. Davis, W. E. Wright, T. Kadesch, C. Murre, A. Voronova, D. Baltimore and H. Weintraub (1991). "Functional-Activity of Myogenic Hlh Proteins Requires Hetero-Oligomerization with E12/E47-Like Proteins In vivo." *Cell* 66: 305-315.
- Lee, W. H., R. Bookstein, F. Hong, L. J. Young, J. Y. Shew and E. Y. Lee (1987). "Human retinoblastoma susceptibility gene: cloning, identification, and sequence." *Science* 235: 1394-9.
- Lee, W. H., J. Y. Shew, F. D. Hong, T. W. Sery, L. A. Donoso, L. J. Young, R. Bookstein and E. Y. Lee (1987). "The retinoblastoma susceptibility gene encodes a nuclear phosphoprotein associated with DNA binding activity." *Nature* 329: 642-5.
- Lochmuller, H., T. Johns and E. A. Shoubridge (1999). "Expression of the E6 and E7 genes of human papillomavirus (HPV16) extends the life span of human myoblasts." *Exp Cell Res* 248: 186-93.
- Ma, P. C., M. A. Rould, H. Weintraub and C. O. Pabo (1994). "Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation." *Cell* 77: 451-9.
- Macleod, K. F., Y. Hu and T. Jacks (1996). "Loss of Rb activates both p53-dependent and independent cell death pathways in the developing mouse nervous system." *EMBO J* 15: 6178-88.
- Magenta, A., C. Cenciarelli, F. De Santa, P. Fuschi, F. Martelli, M. Caruso and A. Felsani (2003). "MyoD stimulates RB promoter activity via the CREB/p300 nuclear transduction pathway." *Mol Cell Biol* 23: 2893-906.
- Markley, J. L., W. J. Horsley and M. P. Klein (1971). "Spin-Lattice Relaxation Measurements in Slowly Relaxing Complex Spectra." *J Chem Phys* 55: 3604-&.

-
- Massari, M. E. and C. Murre (2000). "Helix-loop-helix proteins: Regulators of transcription in eucaryotic organisms." *Mol Cell Biol* 20: 429-440.
- McLachlan, A. D. (1979). "Gene duplications in the structural evolution of chymotrypsin." *J Mol Biol* 128: 49-79.
- Miesenbock, G., D. A. De Angelis and J. E. Rothman (1998). "Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins." *Nature* 394: 192-5.
- Minks, C., R. Huber, L. Moroder and N. Budisa (2000). "Noninvasive tracing of recombinant proteins with "fluorophenylalanine-fingers"." *Anal Biochem* 284: 29-34.
- Mok, Y. K., E. L. Elisseeva, A. R. Davidson and J. D. Forman-Kay (2001). "Dramatic stabilization of an SH3 domain by a single substitution: roles of the folded and unfolded states." *J Mol Biol* 307: 913-28.
- Moncrieffe, M. C., N. Juranic, M. D. Kemple, J. D. Potter, S. Macura and F. G. Prendergast (2000). "Structure-fluorescence correlations in a single tryptophan mutant of carp parvalbumin: Solution structure, backbone and side-chain dynamics." *J Mol Biol* 297: 147-163.
- Montelione, G. T., D. Zheng, Y. J. Huang, K. C. Gunsalus and T. Szyperski (2000). "Protein NMR spectroscopy in structural genomics." *Nat Struct Biol* 7 Suppl: 982-5.
- Moreno, V., F. X. Bosch, N. Munoz, C. J. Meijer, K. V. Shah, J. M. Walboomers, R. Herrero and S. Franceschi (2002). "Effect of oral contraceptives on risk of cervical cancer in women with human papillomavirus infection: the IARC multicentric case-control study." *Lancet* 359: 1085-92.
- Morgan, D. O. (1997). "Cyclin-dependent kinases: engines, clocks, and microprocessors." *Annu Rev Cell Dev Biol* 13: 261-91.
- Morgenbesser, S. D., B. O. Williams, T. Jacks and R. A. DePinho (1994). "p53-dependent apoptosis produced by Rb-deficiency in the developing mouse lens." *Nature* 371: 72-4.
- Mori, S., C. Abeygunawardana, M. O. Johnson and P. C. van Zijl (1995). "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: 94-8.

-
- Moroni, M. C., E. S. Hickman, E. L. Denchi, G. Caprara, E. Colli, F. Cecconi, H. Muller and K. Helin (2001). "Apaf-1 is a transcriptional target for E2F and p53." *Nat Cell Biol* 3: 552-8.
- Morozov, A., P. Shiyanov, E. Barr, J. M. Leiden and P. Raychaudhuri (1997). "Accumulation of human papillomavirus type 16 E7 protein bypasses G1 arrest induced by serum deprivation and by the cell cycle inhibitor p21." *J Virol* 71: 3451-7.
- Muhlhahn, P., M. Zweckstetter, J. Georgescu, C. Ciosto, C. Renner, M. Lanzendorfer, K. Lang, D. Ambrosius, M. Baier, R. Kurth and T. A. Holak (1998). "Structure of interleukin 16 resembles a PDZ domain with an occluded peptide binding site." *Nat Struct Biol* 5: 682-6.
- Murre, C., P. S. Mccaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub and D. Baltimore (1989). "Interactions between Heterologous Helix-Loop-Helix Proteins Generate Complexes That Bind Specifically to a Common DNA-Sequence." *Cell* 58: 537-544.
- Murzin, A. G., A. M. Lesk and C. Chothia (1994). "Principles Determining the Structure of Beta-Sheet Barrels in Proteins .1. A Theoretical-Analysis." *J Mol Biol* 236: 1369-1381.
- Murzin, A. G., A. M. Lesk and C. Chothia (1994). "Principles Determining the Structure of Beta-Sheet Barrels in Proteins .2. The Observed Structures." *J Mol Biol* 236: 1382-1400.
- Neff, M. M., C. Fankhauser and J. Chory (2000). "Light: an indicator of time and place." *Genes Dev* 14: 257-71.
- Nemetz, C., R. Reichhuber, R. Schweizer, P. Hloch and M. Watzel (2001). "Reliable quantification of in vitro synthesized green fluorescent protein: comparison of fluorescence activity and total protein levels." *Electrophoresis* 22: 966-9.
- Neuhaus, D., Williamson, MP. (2000). *The nuclear overhauser effect in structural and conformational analysis*, 2nd ed. New York, Wiley-VCH.
- Novitch, B. G., G. J. Mulligan, T. Jacks and A. B. Lassar (1996). "Skeletal muscle cells lacking the retinoblastoma protein display defects in muscle gene expression and accumulate in S and G2 phases of the cell cycle." *J Cell Biol* 135: 441-56.

-
- Pagliuca, A., P. C. Bartoli, S. Saccone, G. Dellavalle and L. Lania (1995). "Molecular-Cloning of Id4, a Novel Dominant-Negative Helix-Loop-Helix Human Gene on Chromosome 6p21.3-P22." *Genomics* 27: 200-203.
- Pearson, G. A. (1987). "Optimization of Gaussian Resolution Enhancement." *J Magn Reson* 74: 541-545.
- Pellecchia, M., D. S. Sem and K. Wuthrich (2002). "NMR in drug discovery." *Nat Rev Drug Discov* 1: 211-9.
- Pervushin, K., V. Gallus and C. Ritter (2001). "Improved TROSY-HNCA experiment with suppression of conformational exchange induced relaxation." *J Biomol NMR* 21: 161-6.
- Philips, G. (1997). "Structure and dynamics of green fluorescent protein." *Curr Opin Struct Biol* 7: 821-827.
- Philips, G. (1998). *Green Fluorescent Protein*. New York, Wiley-Liss.
- Poleskaya, A., I. Naguibneva, L. Fritsch, A. Duquet, S. Ait-Si-Ali, P. Robin, A. Vervisch, L. L. Pritchard, P. Cole and A. Harel-Bellan (2001). "CBP/p300 and muscle differentiation: no HAT, no muscle." *EMBO J* 20: 6816-25.
- Prabhu, S., A. Ignatova, S. T. Park and X. H. Sun (1997). "Regulation of the expression of cyclin-dependent kinase inhibitor p21 by E2A and Id proteins." *Mol Cell Biol* 17: 5888-96.
- Prendergast, F. G. (1999). "Biophysics of the green fluorescent protein." *Methods Cell Biol* 58: 1-18.
- Prestegard, J. H., H. Valafar, J. Glushka and F. Tian (2001). "Nuclear magnetic resonance in the era of structural genomics." *Biochemistry* 40: 8677-85.
- Pucci, B., M. Kasten and A. Giordano (2000). "Cell cycle and apoptosis." *Neoplasia* 2: 291-9.
- Puri, P. L., M. L. Avantaggiati, C. Balsano, N. Sang, A. Graessmann, A. Giordano and M. Levvero (1997). "p300 is required for MyoD-dependent cell cycle arrest and muscle-specific gene transcription." *EMBO J* 16: 369-83.
- Puri, P. L., S. Iezzi, P. Stiegler, T. T. Chen, R. L. Schiltz, G. E. Muscat, A. Giordano, L. Kedes, J. Y. Wang and V. Sartorelli (2001). "Class I histone deacetylases sequentially interact with MyoD and pRb during skeletal myogenesis." *Mol Cell* 8: 885-97.

-
- Reif, B., M. Hennig and C. Griesinger (1997). "Direct measurement of angles between bond vectors in high-resolution NMR." *Science* 276: 1230-3.
- Renan, M. J. (1993). "How Many Mutations Are Required for Tumorigenesis - Implications from Human Cancer Data." *Mol Carcinog* 7: 139-146.
- Reuter, W. and C. Nickelreuter (1993). "Molecular Assembly of the Phycobilisomes from the Cyanobacterium *Mastigocladus-Laminosus*." *J Photochem Photobiol B Biol* 18: 51-66.
- Reynaud, E. G., K. Pelpel, M. Guillier, M. P. Leibovitch and S. A. Leibovitch (1999). "p57(Kip2) stabilizes the MyoD protein by inhibiting cyclin E-Cdk2 kinase activity in growing myoblasts." *Mol Cell Biol* 19: 7621-9.
- Riechmann, V. and F. Sablitzky (1995). "Mutually Exclusive Expression of 2 Dominant-Negative Helix-Loop-Helix (Dnhlh) Genes, Id4 and Id3, in the Developing Brain of the Mouse Suggests Distinct Regulatory Roles of These Dnhlh Proteins During Cellular Proliferation and Differentiation of the Nervous-System." *Cell Growth Differ* 6: 837-843.
- Riek, R., K. Pervushin and K. Wuthrich (2000). "TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution." *Trends Biochem Sci* 25: 462-8.
- Sambrook, J., Russell, DW. (2001). *Molecular Cloning*, CSHL press, Cold Spring Harbor.
- Sandal, T. (2002). "Molecular aspects of the mammalian cell cycle and cancer." *Oncologist* 7: 73-81.
- Santos, C., N. Munoz, S. Klug, M. Almonte, I. Guerrero, M. Alvarez, C. Velarde, O. Galdos, M. Castillo, J. Walboomers, C. Meijer and E. Caceres (2001). "HPV types and cofactors causing cervical cancer in Peru." *Br J Cancer* 85: 966-71.
- Sawasaki, T., Y. Hasegawa, M. Tsuchimochi, N. Kamura, T. Ogasawara, T. Kuroita and Y. Endo (2002). "A bilayer cell-free protein synthesis system for high-throughput screening of gene products." *FEBS Lett* 514: 102-5.
- Schagger, H. and G. Vonjagow (1987). "Tricine Sodium Dodecyl-Sulfate Polyacrylamide-Gel Electrophoresis for the Separation of Proteins in the Range from 1-Kda to 100-Kda." *Anal Biochem* 166: 368-379.
- Schatz, G. W., J. Reinking, J. Zippin, L. K. Nicholson and V. M. Vogt (2001). "Importance of the N terminus of rous sarcoma virus protease for structure and enzymatic function." *J Virol* 75: 4761-70.

-
- Schneider, J. W., W. Gu, L. Zhu, V. Mahdavi and B. Nadal-Ginard (1994). "Reversal of terminal differentiation mediated by p107 in Rb^{-/-} muscle cells." *Science* 264: 1467-71.
- Schwede, T., J. Kopp, N. Guex and M. C. Peitsch (2003). "SWISS-MODEL: An automated protein homology-modeling server." *Nucleic Acids Res* 31: 3381-5.
- Seifert, M. H., J. Georgescu, D. Ksiazek, P. Smialowski, T. Rehm, B. Steipe and T. A. Holak (2003). "Backbone dynamics of green fluorescent protein and the effect of histidine 148 substitution." *Biochemistry* 42: 2500-12.
- Seifert, M. H., D. Ksiazek, M. K. Azim, P. Smialowski, N. Budisa and T. A. Holak (2002). "Slow exchange in the chromophore of a green fluorescent protein variant." *J Am Chem Soc* 124: 7932-42.
- Serrano, M., G. J. Hannon and D. Beach (1993). "A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4." *Nature* 366: 704-7.
- Sheikh, A. and R. B. Freedman (1990). "Translation of preprochymosin in vitro. Evidence for folding of prochymosin to the native conformation." *Biochem J* 272: 659-64.
- Shen, C. P. and T. Kadesch (1995). "B-Cell-Specific DNA-Binding by an E47 Homodimer." *Mol Cell Biol* 15: 4518-4524.
- Sherr, C. J. (1996). "Cancer cell cycles." *Science* 274: 1672-1677.
- Sherr, C. J. (2000). "The Pezcoller Lecture: Cancer cell cycles revisited." *Cancer Res* 60: 3689-3695.
- Sherr, C. J. and J. M. Roberts (1999). "CDK inhibitors: positive and negative regulators of G(1)-phase progression." *Genes Dev* 13: 1501-1512.
- Shuker, S. B., P. J. Hajduk, R. P. Meadows and S. W. Fesik (1996). "Discovering high-affinity ligands for proteins: SAR by NMR." *Science* 274: 1531-4.
- Skelton, N. J., A. G. Palmer, M. Akke, J. Kordel, M. Rance and W. J. Chazin (1993). "Practical Aspects of 2-Dimensional Proton-Detected N-15 Spin Relaxation Measurements." *J Magn Reson B* 102: 253-264.
- Sklenar, V., M. Piotto, R. Leppik and V. Saudek (1993). "Gradient-Tailored Water Suppression for H-1-N-15 Hs qc Experiments Optimized to Retain Full Sensitivity." *J Magn Reson A* 102: 241-245.
- Sklenar, V., D. Torchia and A. Bax (1987). "Measurement of C-13 Longitudinal Relaxation Using H-1 Detection." *J Magn Reson* 73: 375-379.

-
- Stoll, R., C. Renner, S. Hansen, S. Palme, C. Klein, A. Belling, W. Zeslawski, M. Kamionka, T. Rehm, P. Muhlhahn, R. Schumacher, F. Hesse, B. Kaluza, W. Voelter, R. A. Engh and T. A. Holak (2001). "Chalcone derivatives antagonize interactions between the human oncoprotein MDM2 and p53." *Biochemistry* 40: 336-44.
- Storf, M., A. Parbel, M. Meyer, B. Strohmam, H. Scheer, M. G. Deng, M. Zheng, M. Zhou and K. H. Zhao (2001). "Chromophore attachment to biliproteins: specificity of PecE/PecF, a lyase-isomerase for the photoactive 3(1)-cys-alpha 84-phycoviolobilin chromophore of phycoerythrocyanin." *Biochemistry* 40: 12444-56.
- Tietze, K., N. Oellers and E. Knust (1992). "Enhancer of Split(D), a Dominant Mutation of *Drosophila*, and Its Use in the Study of Functional Domains of a Helix Loop Helix Protein." *Proc Natl Acad Sci U S A* 89: 6152-6156.
- Tjandra, N., J. G. Omichinski, A. M. Gronenborn, G. M. Clore and A. Bax (1997). "Use of dipolar ¹H-¹⁵N and ¹H-¹³C couplings in the structure determination of magnetically oriented macromolecules in solution." *Nat Struct Biol* 4: 732-8.
- Tjandra, N., A. Szabo and A. Bax (1996). "Protein backbone dynamics and N-15 chemical shift anisotropy from quantitative measurement of relaxation interference effects." *J Am Chem Soc* 118: 6986-6991.
- Toby, G. G. and E. A. Golemis (2001). "Using the yeast interaction trap and other two-hybrid-based approaches to study protein-protein interactions." *Methods* 24: 201-17.
- Tsien, R. Y. (1998). "The green fluorescent protein." *Annu Rev Biochem* 67: 509-44.
- Van Bramer, S. (1999). "An introduction to the Fourier transform." *J Chem Educ* 76.
- van Thor, J. J., T. Gensch, K. J. Hellingwerf and L. N. Johnson (2002). "Phototransformation of green fluorescent protein with UV and visible light leads to decarboxylation of glutamate 222." *Nat Struct Biol* 9: 37-41.
- van Thor, J. J., A. J. Pierik, I. Nugteren-Roodzant, A. Xie and K. J. Hellingwerf (1998). "Characterization of the photoconversion of green fluorescent protein with FTIR spectroscopy." *Biochemistry* 37: 16915-21.
- Voityuk, A. A., M. E. MichelBeyerle and N. Rosch (1997). "Protonation effects on the chromophore of green fluorescent protein. Quantum chemical study of the absorption spectrum." *Chem Phys Lett* 272: 162-167.
- Voityuk, A. A., M. E. Michel-Beyerle and N. Rosch (1998). "Structure and rotation barriers for ground and excited states of the isolated chromophore of the green fluorescent protein." *Chem Phys Lett* 296: 269-276.

- Wachter, R. M., M. A. Elsliger, K. Kallio, G. T. Hanson and S. J. Remington (1998). "Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein." *Structure* 6: 1267-77.
- Wang, J. and M. F. Wilkinson (2001). "Deletion mutagenesis of large (12-kb) plasmids by a one-step PCR protocol." *Biotechniques* 31: 722-4.
- Wang, W. and B. A. Malcolm (1999). "Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis." *Biotechniques* 26: 680-2.
- Ward, W. (1981). *Bioluminescence and chemiluminescence: Basic Chemistry and Analytical Applications*, Academic Press, New York.
- Weber, W., V. Helms, J. A. McCammon and P. W. Langhoff (1999). "Shedding light on the dark and weakly fluorescent states of green fluorescent proteins." *Proc Natl Acad Sci U S A* 96: 6177-6182.
- Weinberg, R. A. (1991). "Tumor suppressor genes." *Science* 254: 1138-46.
- Weintraub, H., Tapscott, S.J., Davis, R.L., Thayer, M.J., Adam, M.A., Lassar, A.B., Miller, A.D. (1989). "Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD." *Proc Natl Acad Sci U S A* 86: 5434-8.
- Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T. K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, Y. Zhuang and A. Lassar (1991). "The MyoD Gene Family - Nodal Point During Specification of the Muscle-Cell Lineage." *Science* 251: 761-766.
- Whyte, P., K. J. Buchkovich, J. M. Horowitz, S. H. Friend, M. Raybuck, R. A. Weinberg and E. Harlow (1988). "Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product." *Nature* 334: 124-9.
- Wiegand, G., A. Parbel, M. H. Seifert, T. A. Holak and W. Reuter (2002). "Purification, crystallization, NMR spectroscopy and biochemical analyses of alpha-phycoerythrocyanin peptides." *Eur J Biochem* 269: 5046-55.
- Wijesinha-Bettoni, R., C. M. Dobson and C. Redfield (2001). "Comparison of the structural and dynamical properties of holo and apo bovine alpha-lactalbumin by NMR spectroscopy." *J Mol Biol* 307: 885-98.
- Wüthrich, K. (1994). *NMR of proteins and nucleic acids*. New York, John Wiley & Sons.

-
- Yang, F., L. G. Moss and G. N. Phillips (1996). "The molecular structure of green fluorescent protein." *Nature Biotechnol* 14: 1246-1251.
- Zabludoff, S. D., M. Csete, R. Wagner, X. Yu and B. J. Wold (1998). "p27Kip1 is expressed transiently in developing myotomes and enhances myogenesis." *Cell Growth Differ* 9: 1-11.
- Zacharias, D. A., J. D. Violin, A. C. Newton and R. Y. Tsien (2002). "Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells." *Science* 296: 913-6.
- Zhang, J. M., Q. Wei, X. Zhao and B. M. Paterson (1999). "Coupling of the cell cycle and myogenesis through the cyclin D1-dependent interaction of MyoD with cdk4." *EMBO J* 18: 926-33.
- Zhang, J. M., X. Zhao, Q. Wei and B. M. Paterson (1999). "Direct inhibition of G(1) cdk kinase activity by MyoD promotes myoblast cell cycle withdrawal and terminal differentiation." *EMBO J* 18: 6983-93.
- Zhao, K. H., R. Haessner, E. Cmiel and H. Scheer (1995). "Type-I Reversible Photochemistry of Phycoerythrocyanin Involves Z/E-Isomerization of Alpha-84 Phycoviolobilin Chromophore." *Biochim. Biophys. Acta* 1228: 235-243.
- Zhao, K. H., J. P. Zhu, M. G. Deng, M. Zhou, M. Storf, A. Parbel and H. Scheer (2003). "Photochromic chromopeptides derived from phycoerythrocyanin: biophysical and biochemical characterization." *Photochem Photobiol Sci* 2: 741-748.
- Zurdo, J., J. I. Guijarro, J. L. Jimenez, H. R. Saibil and C. M. Dobson (2001). "Dependence on solution conditions of aggregation and amyloid formation by an SH3 domain." *J Mol Biol* 311: 325-40.
- Zweckstetter, M. and T. A. Holak (1998). "An adiabatic multiple spin-echo pulse sequence: removal of systematic errors due to pulse imperfections and off-resonance effects." *J Magn Reson* 133: 134-47.

7 SUPPLEMENTARY MATERIALS

7.1 CDK4/6 structurally guided alignment

TARGET	1		MATSR	YECVAEIGEG	AYGTVYKARD	-HSGHFVA	LKSVRVENGG
lblxA	5	SLC--	RADQQ	YECVAEIGEG	AYGKVFVKARD	LKNGGRFVA	LKRVRVQTGE
lg3nE	9		ADQQ	YECVAEIGEG	AYGKVFVKARD	LKNGGRFVA	LKRVRVQTGE
lg3nA	9		ADQQ	YECVAEIGEG	AYGKVFVKARD	LKNGGRFVA	LKRVRVQTGE
ljowB	10		DOQ	YECV	YGKVFVKARD	LKNGGRFVA	LKRVRVQTGEE
lbi7A	10	DC--Q	YECVAEIGEG	AYGKVFVKARD	LKNGGRFVA	LKRVRV----	

			** * .	** *****	..* ***	** .***	
TARGET			ssssssssss	ssssssss	ssss	ssssssssss	
lblxA		hhs	ssssssssss	ssssssss	s	ssss	ssssssssss
lg3nE		s	ssssssssss	ssssssss	s	ssss	ssssssssss
lg3nA		s	ssssssssss	ssssssss	s	ssss	ssssssssss
ljowB			ss	sss	sssss	sss	
lbi7A			ssssssssss	ssssssss	ssss	ssss	

TARGET	44	GCGGG-LHIS	TVREVALLR	LEAFEHPNVV	RLMDVCATSR	TDREIKVTLV
lblxA	52	EG----MFLS	TIREVAVLRH	LETFEHPNVV	RLFDVCTVSR	TDRETKLTLV
lg3nE	52	EG----MFLS	TIREVAVLRH	LETFEHPNVV	RLFDVCTVSR	TDRETKLTLV
lg3nA	52	EG----MFLS	TIREVAVLRH	LETFEHPNVV	RLFDVCTVSR	TDRETKLTLV
ljowB	53	G----MFLS	TIREVAVLRH	LETFEHPNVV	RLFDVCTVSR	TDRETKLTLV
lbi7A	48	-----	-----	-----	RLFDVCTVSR	TDRETKLTLV

			sssh	hhhhhhhhh	hhh	ssssssss	ssssssss
TARGET			sssh	hhhhhhhhh	hhh	ssssssssss	s sssssss
lblxA			hh	hhhhhhhhh	hhh	ssssssssss	ssssssss
lg3nE			hh	hhhhhhhhh	hhh	ssssssssss	ssssssss
lg3nA			hh	hhhhhhhhh	hhh	ssssssssss	ssssssss
ljowB				hhhhhhh	h	ssssssss	ssss
lbi7A						ssssssssss	s sssssss

TARGET	93	FEHVDQDLRT	YLDKAPFPGL	PAETIKDLMR	QFLRGLDFLH	ANCIVHRDLK
lblxA	98	FEHVDQDLTT	YLDKVEPEGV	PTETIKDMMF	QLLRGLDFLH	SHRVVHRDLK
lg3nE	98	FEHVDQDLTT	YLDKVEPEGV	PTETIKDMMF	QLLRGLDFLH	SHRVVHRDLK
lg3nA	98	FEHVDQDLTT	YLDKVEPEGV	PTETIKDMMF	QLLRGLDFLH	SHRVVHRDLK
ljowB	98	FEHVDQDLTT	YLDKVEPEGV	PTETIKDMMF	QLLRGLDFLH	SHRVVHRDLK
lbi7A	98	FEHVDQDLTT	YLDKVEPEGV	PTETIKDMMF	QLLRGLDFLH	SHRVVHRDLK

			ss	sssh	hhhh	hhhhhhhhh	hhhhhhhhh
TARGET			ss	sssh	hhhh	hhhhhhhhh	hhhhhhhhh h
lblxA			ss	sssh	hh	hhhhhhhhh	hhhhhhhhh
lg3nE			ss	sssh	hh	hhhhhhhhh	hhhhhhhhh
lg3nA			ss	sssh	hh	hhhhhhhhh	hhhhhhhhh
ljowB			sss	sssh	hhhh	hhhhhhhhh	hhhhhhhhh sss
lbi7A			ss	sssh	hhh	hhhhhhhhh	hhhhhhhhh

TARGET	143	PENILVTSSG	TVKLADFGLA	RIYSYQMALT	SVVVTWLYR	APEVLLQSTY
lblxA	148	PQNILVTSSG	QIKLADFGLA	RIYSFQMALT	SVVVTWLYR	APEVLLQSSY
lg3nE	148	PQNILVTSSG	QIKLADFGLA	RIYSFQMALT	SVVVTWLYR	APEVLLQSSY
lg3nA	148	PQNILVTSSG	QIKLADFGLA	RIYSFQMALT	SVVVTWLYR	APEVLLQSSY
ljowB	148	PQNILVTSSG	QIKLADFGLA	RIYSFQMALT	SVVVTWLYR	APEVLLQSSY
lbi7A	148	PQNILVTSSG	QIKLADFGLA	RIYSFQMALT	SVVVTWLYR	APEVLLQSSY

			* .***** *	*****	****	*****	*****
TARGET			sssss	sssss		hhhhh	
lblxA			sssss	sssss		hhhh	
lg3nE			sssss	sssss	sssss	s ssssh	hhhhh
lg3nA			sssss	sssss	sssss	s ssssh	hhhhh
ljowB			sssss	sssss	sss		hhhh

1bi7A		sssss	sssss	sss	sss	hhhhh
TARGET	192	A T P V D M W S V G	C I F A E M F R R K	F L F C G N S E A D	Q L G K I F D L I G	L P P E D D W E R D
1blxA	197	A T P V D L W S V G	C I F A E M F R R K	F L F R G S S D V D	Q L G K I L D V I G	L P G E E D W E R D
1g3nE	197	A T P V D L W S V G	C I F A E M F R R K	F L F R G S S D V D	Q L G K I L D V I G	L P G E E D W E R D
1g3nA	197	A T P V D L W S V G	C I F A E M F R R K	F L F R G S S D V D	Q L G K I L D V I G	L P G E E D W E R D
1jowB	197	A T P V D L W S V G	C I F A E M F R R K	F L F R G S S D V D	Q L G K I L D V I G	L P G E E D W E R D
1bi7A	197	A T P V D L W S V G	C I F A E M F R R K	F L F R G S S D V D	Q L G K I L D V I G	L P G E E D W E R D
		*****	*****	***	*.*.*	*****
TARGET		hhhhhhhh	hhhhhhhh	hhh	hhhhhhhhh	
1blxA		hhhhhhhh	hhhhhhhh	hhh	hhhhhhhhh	
1g3nE		hhhhhhhh	hhhhhhhh	hhh	hhhhhhhhh	
1g3nA		hhhhhhhh	hhhhhhhh	hhh	hhhhhhhhh	
1jowB		hhhhhhhh	hhhh	hhh	hhhhhhhhh	
1bi7A		hhhhhhhh	hhhhhhhh	hhh	hhhhhhhhh	
TARGET	242	V S L E R Q A F H S	R G I R E V Q S V V	D L M E E S G A Q L	L L E M L T F N E A	K R I S A F R A L Q
1blxA	247	V A L P R Q A F H S	K S A Q E I E K F V	T D I D E L G K D L	L L K C L T F N E A	K R I S A Y S A L S
1g3nE	247	V A L P R Q A F H S	K S A Q E I E K F V	T D I D E L G K D L	L L K C L T F N E A	K R I S A Y S A L S
1g3nA	247	V A L P R Q A F H S	K S A Q E I E K F V	T D I D E L G K D L	L L K C L T F N E A	K R I S A Y S A L S
1jowB	252	-----QAFHS	KSAQEIEKFFV	TDIDELGKDL	LLKCLTFNEA	KRISAYSALS
1bi7A	247	V A L P R Q A F H S	K S A Q E I E K F V	T D I D E L G K D L	L L K C L T F N E A	K R I S A Y S A L S
		**	*.*.*.*	***	*.*.*	*****
TARGET				hhhhhh	hhhh	hhhhhh
1blxA				hhhhhh	hhhh	hhhhhh
1g3nE				hhhhhh	hhhh	hhhhhh
1g3nA				hhhhhh	hhhh	hhhhhh
1jowB			hhh	hhhhhh	h	hhhhhh
1bi7A				hhhhhh	hhhh	hhhhhh
TARGET	292	H S Y L H K D E G N	--			
1blxA	297	H P Y F Q L E R C	K E N			
1g3nE	297	H P Y F C	---			
1g3nA	297	H P Y F C	---			
1jowB	297	H P Y F	---			
1bi7A	297	H P Y F C	---			
		.				
TARGET		hh				
1blxA		hh				
1g3nE						
1g3nA						
1jowB						
1bi7A						

Structural guided alignment of CDK4 over CDK6 sequences taken from PDB files (Guex et al. 1997; Schwede et al. 2003).

- 1blxA p19/CDK6 resolution
- 1g3nE p18/CDK6/cycline K resolution
- 1g3nA p18/CDK6/cycline K resolution
- 1jowB CDK6/Viral cycline resolution 3.1
- 1bi7A CDK6/p16 resolution 3.4

TARGET CDK4 sequence
 CDK4/6 share 71.1% identity in 294 residues overlap; Score: 1063.0; Gap frequency: 1.4%

Color code of residues:

Polar positive	H, K, R	Blue
Polar negative	D, E	Red
Polar neutral	S, T, N, Q	Green
Non-polar aliphatic	A, V, L, I, M	White
Non-polar aromatic	F, Y, W	Purple
	P, G	Brown
	C	Yellow
Special characters	B, Z, X, -	Grey

Symbols

* identical residue
 s strand
 h helix

7.2 Bioinformatics und calculations

List of programs and web-pages used to analyse sequences and, structures:

Nucleic acid

Single sequence

- Similarity search [BLAST](#), [BLAST 2](#)
- Feature search
 - Restriction sites [BCM](#), [Webcutter](#), [Molecular Toolkit](#)
 - ORF, start, stop, poly(A)-signalling sites [GeneScan](#), [ORF finder](#), [WebGene](#)
 - Splice sites [SSP by NN](#)
 - Transcription factor binding sites [TF Search](#)
- PCR primer design [BCM](#), [Primer3](#), [NetPrimer](#), [SDM primer generator](#)
- Search for orthologs and homologs [BLAST](#), [EST at NCBI](#)

- Translation [BCM](#)
- Codon usage [Alces](#), [Codon Usage Database](#), [E.coli codon usage](#), [E.coli codon usage analyzer](#)

Multiple sequences

- Aligment [BCM](#)
- Generation of consensus sequence [BCM](#), [BLAST](#)
- Phylogenetic analisys [Philip](#), [Phylogendron](#)

Proteins

Single sequence

- Similarity search [BLAST](#), [PropSearch](#)
- Features
 - MW, pI, AA composition [ProtParam at ExPaSy](#)
 - Hydrophobicity [BCM](#)
 - Transmembrane region [TMPred](#)
 - Post-translational modification sites
 - O- glycosylation [NetOGlyc](#)
 - Phosphorylation [NetPhos](#)
 - Signal sequences [Psort](#)
- Secondary structure predictions [PredictProtein](#), [nnPredict](#)
- Motifs [MAST](#), [PRINTS](#) [BLAST](#)
- Molecular modelling [SWISS-MODEL](#)
- Reverse translation [Molecular Toolkit](#)
- Codon usage [Alces](#), [Codon Usage Database](#), [E.coli codon usage](#), [E.coli codon usage analyzer](#)

Multiple sequences

- Aligment [BCM](#)
- Phylogenetic analisys [Philip](#), [Phylogendron](#)

Expression system

- Codon usage [Alces](#), [Codon Usage Database](#), [E.coli codon usage](#), [E.coli codon usage analyzer](#)
- Vectors databases [VAST at NCBI](#)
- Strains databases [ATCC](#), [E.coli INDEX](#)

Protocols

- [Protocol-Online](#), [Highveld](#), [NWFSC Molecular Biology Protocols](#), [Gerard R. Lazo](#)

NMR data analysis

- Assigment Sparky (Goddard 2001)
- In-house software CCNMR (Cieslar et al. 1990)

ACKNOWLEDGEMENT

I would like to thank everyone who contributed to this work. In particular to Professor Rober Huber for being my Doktorvater, and, for the opportunity of working in his department.

To my supervisor Doctor Tad A. Holak for his many scientific contributions, guidance and interest in my work.

To the NMR master, Markus Seifert for showing me the power of the Nuclear Magnetic Resonance.

To the NMR and MPI team (Loyola D'Silva, Madhumita Ghosh, Marcin Krajewski, Joma Kanikadu Joy, Sudipta Majumdar, Narashimha Rao Nalabothula, Przemyslaw Ozdowy, Grzegorz Popwicz, Ulli Rothweiler, Mahavir Singh, Igor Siwanowicz, Ashwani Thakur, Christian Renner, Julia Georgescu, Raphael Stoll, Conny Ciosto, Mariusz Kamionka, Sreejesh Shanker, Markus Seifert, Michael Brüggert, Dorota Ksiazek, André Schönichen, Till Rehm, Chrystelle Mavoungou) for their help, advice and good working atmosphere.

Last but not least, I would like to acknowledge my family and Ola for all the love, care and support.