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

Structural Investigations on Proteins Involved in Cancer Development

Mariusz Kamionka

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. Wolfgang Hiller
Prüfer der Dissertation:	1.	apl.-Prof. Dr. Luis Moroder
	2.	Univ.-Prof. Dr. Johannes Buchner

Die Dissertation wurde am 5.09.2001 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 15.10.2001 angenommen.

Acknowledgements

I would like to thank all of those who have contributed to this work.

In particular I am most grateful to Professor Robert Huber for giving me the opportunity to work in his department and for access to all his laboratories and facilities, and to Professor Luis Moroder for being my Doktorvater.

This thesis was only possible because of the support of Doctor Tad A. Holak, my supervisor, to whom I am indebted not just for his scientific contribution but also for his motivating words, day after day, his help, financial support and his friendship.

My thank goes to all of the NMR friendly team for their help and advice, physicists: Dr. Ruth Pfänder, Till Rehm, Markus Seifert, as well as people who worked with me in the lab: Anja Belling, Michael Brüggert, Dr. Cornelia Ciosto, Dr. Julia Georgescu, Madhumita Ghosh, Chrystelle Mavoungou, Dr. Peter Mühlhahn, Narashimha Rao Nalabothula, Sreejesh Shanker, Igor Siwanowicz and Paweł Śmiałowski.

In particular I am grateful to Dorota Książek and Dr. Wojciech Żesławski for their scientific but also emotional support.

My apologies to the others who I have not mentioned by name, I am indebted to them for the many ways they helped me.

Finally, I would like to pay tribute to the constant support of my family and my friends, without their love over the many months none of this would have been possible and whose sacrifice I can never repay.

The Lord had said unto Abram, Get thee out of thy country, and from thy kindred, and from thy father's house, unto a land that I will show thee:
And I will make of thee a great nation, and I will bless thee, and make thy name great.

(Genesis 12:1.2)

Publications

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

Raphael Stoll, Christian Renner, Silke Hansen, Stefan Palme, Christian Klein, Anja Belling, Wojciech Zeslawski, **Mariusz Kamionka**, Till Rehm, Peter Mühlhahn, Ralf Schumacher, Friederike Hesse, Brigitte Kaluza, Wolfgang Voelter, Richard A. Engh and Tad A. Holak
Chalcone Derivatives Antagonize Interactions between the Human Oncoprotein MDM2 and p53
Biochemistry (2001) **40**, 336-344

Wojciech Zeslawski, Hans-Georg Beisel, **Mariusz Kamionka**, Wenzel Kalus, Richard A. Engh, Robert Huber, Kurt Lang and Tad A. Holak
The Interaction of Insulin-like Growth Factor-I with the N-terminal Domain of IGFBP-5
EMBO J. (2001) **20**, 3638-3644

Mariusz Kamionka, Till Rehm, Richard A. Engh, Kurt Lang and Tad A. Holak
A designed small tyrosine derivative that inhibits interaction between IGF-I and IGFBP-5
submitted

Mariusz Kamionka, Dorota Ksiazek, Till Rehm, Chrystelle Mavoungou, Lars Israel, Michael Schleicher and Tad A. Holak
Structure of the N-terminal Domain of Cyclase-Associated Protein (CAP) from *Dictyostelium Discoideum.*
manuscript in preparation

Contents

1 Introduction

Cancer, being one of the major human health problems, has received enormous biomedical attention over past few decades. Malignant tumors are responsible for around 20% mortalities in developed countries. Around 100 different types of human cancers are recognized. We are, however, just beginning to understand the biochemical basis of this collection of diseases. Understanding the cancer development on the molecular level seems to be the most rational way on which in consequence medicines for the disease can be proposed.

Every cell in human organism remains under strict developmental control. Most adult body cells are quiescent. They can, however, lose their developmental controls and commence excessive proliferation, the first step in cancerogenesis, when, for example, an extracellular regulator of the cell development becomes deregulated. This is the case for the tumors connected with deregulation of insulin-like growth factors (IGFs). The structure solution of the IGF-I complex with its regulator IGF-binding protein-5 (IGFBP-5) (chapter 5.3) is the first step towards rational drug design for all diseases in which IGF-I remains out of control.

Eukaryotic cells have many defense mechanisms against cancerogenetic transformation. For example tumor suppressor proteins, which act as inhibitors of tumor development. The best investigated and known example is p53. Deregulation of p53, either by mutations, or by inhibition by other proteins was found in majority of human cancers. Murine Double Minute Clone 2 protein (MDM2) is responsible for keeping p53 level under control. Interaction of these two proteins leads to p53 inhibition and consequent degradation. Excessive expression of MDM2 is very often observed in many types of cancer. Hence, disruption of p53/MDM2 complex might be effective in cancer therapy. Characterizing the binding mode of small organic compounds acting as antagonists of these interaction, might offer the basis for structure-based drug design (chapter 5.2).

Retinoblastoma susceptibility protein (RB) is another well investigated tumor suppressor. It plays an essential role in cell cycle regulation. Proteins involved in cell cycle control are very often found mutated in cancers. In fact, it is true that disabling the RB and p53 pathways is a hallmark of cancer. RB functions as inhibitor of E2F-activity dependent genes that are consequently involved in cell proliferation. Thus, there is growing interest to

explain the manner of E2F function, in particular its interaction with RB tumor suppressor. Chapter 5.1 presents some preliminary investigations of E2F protein family.

Cancer is a collection of many diseases. It explains why searching for a cure for it was performed on many different levels. Additionally, the work presented here is original by employing many various methods. For structure determination x-ray crystallography was used. NMR was employed to search for small organic compounds as potential inhibitors of proteins interaction. Most of the work, described in this thesis, however, was done in laboratory, using biochemistry as well as molecular biology methods.

2 Biological Background

2.1 Cancer Cell Cycle

The cell cycle, the general sequence of events that occur during the lifetime of a eukaryotic cell, is divided into four distinct phases: **M**itosis and cell division occur during the relatively brief **M phase** (around 1h). This is followed by the **G_1 phase** (for **g**ap, around 10h), which covers the longest part of the cell cycle. G_1 gives way to **S phase** (for **s**ynthesis, 6-8h), which, in contrast to events in prokaryotes, is the only period in the cell cycle when DNA is synthesized. During the relatively short **G_2 phase** (2-6h), the now tetraploid cell prepares for mitosis. It then enters M phase once again and thereby commences a new round of the cell cycle (Voet D. & Voet J. G., 1995). Late in the G_1 phase cell makes "decision" whether to proceed to S phase or to withdraw into a resting state (G_0 phase). This is known as the restriction point. One of the tasks of the restriction point is to ensure that DNA is not damaged. Too many mutations, deletions etc. could increase the risk of cancer or other cell dysfunctions.

G_1 to S phase progression begins with cell stimulation by mitogens like EGF (**e**pidermal **g**rowth **f**actor) or CSF-1 (**c**olony-**s**timulating **f**actor 1) (Matsushime et al., 1991). With the time cell becomes refractory to extracellular growth regulatory signals and commit to the autonomous program that carries it through to division. Passage through the restriction point is controlled by **c**yclin **d**ependent **k**inases (CDKs), which are regulated by cyclins D, E and A (Figure 2.1). For its action CDK requires cyclin binding and phosphorylation by **C**DK **a**ctivating **k**inase (CAK). Stimulation of the cell by mitogens effects a signal transduction pathway, which in consequence leads to cyclin D activation. Thus, proteins of cyclin D family (cyclin D1, D2 and D3) act as growth factor sensors. Their catalytic partners are CDK4 and CDK6. Specific polypeptide **in**hibitors of CD**K4** and CDK6 –so called INK4 proteins –can directly block cyclin D dependent kinase activity and cause G_1 phase arrest. There are four known members of INK4 protein family: $p16^{INK4a}$, $p15^{INK4b}$, $p18^{INK4c}$, $p19^{INK4d}$. They can specifically bind and inhibit CDK4 and CDK6, but not other CDKs. Activated cyclin D-dependent kinases are able to phosphorylate retinoblastoma susceptibility protein (RB) and other RB-like proteins (p130, p107), so called "pocket proteins". RB protein controls in turn gene expression mediated by a family of heterodimeric transcriptional regulators called here for simplicity E2Fs (for details see chapter 2.2). It is phosphorylation of RB by CDKs that causes E2Fs release, enabling them to activate genes, which are important for S phase entry

(e.g. **di**hydro**f**olate **r**eductase -DHFR, **t**hymidine **k**inase -TK, **t**hymidylate **s**ynthase -TS, DNA polymerase α -POL, cyclin dependent kinase 1 -CDC2, cyclin E and A, E2F-1 itself).

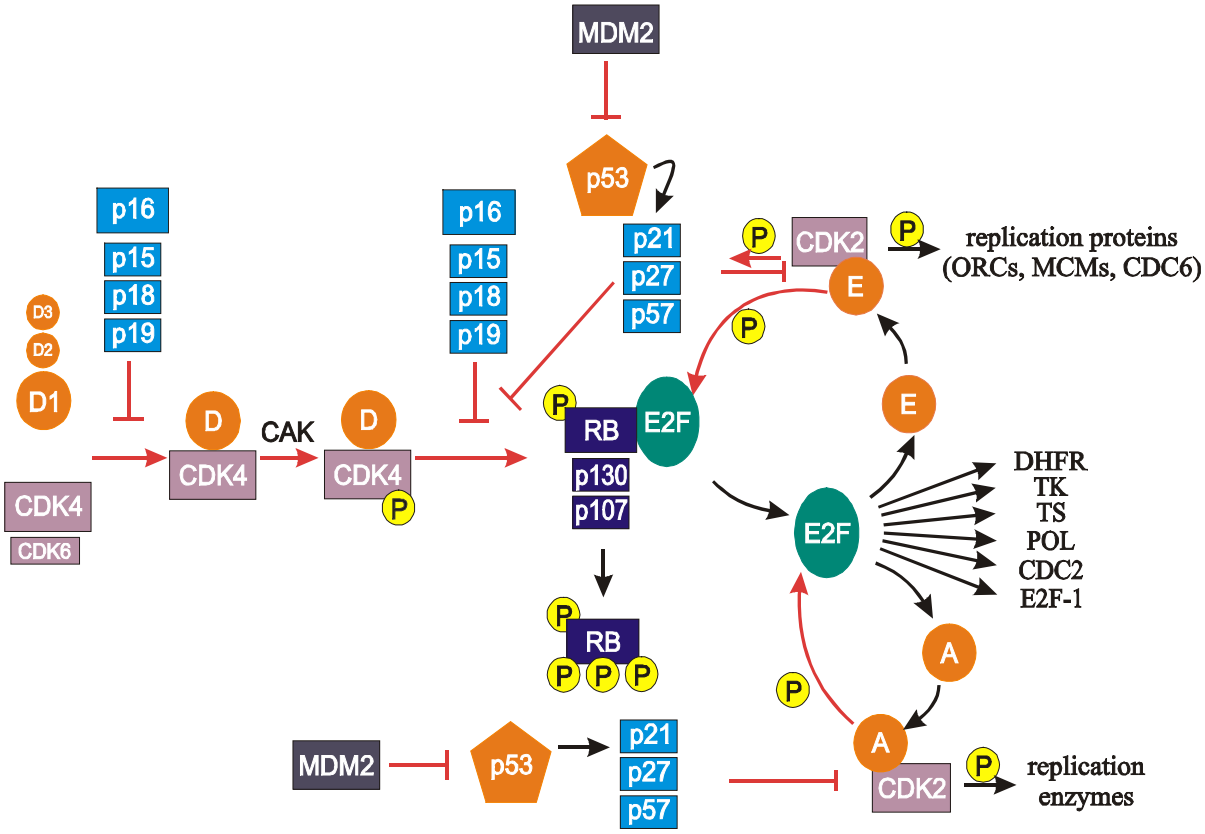


Figure 2.1. *Restriction point control. The proteins most frequently targeted in human cancers are highlighted. CDK inhibitors are shown in light blue, cyclins in orange, CDKs in purple. Arrows depicting inhibitory phosphorylation (P) or inactivating steps are shown in red, and those depicting activating steps are shown in black (Sherr et al., 1996). For details refer to text.*

Expression of cyclin E makes entry into S phase independent on extracellular stimulators. Cyclin E in complex with CDK2 is able to finish RB phosphorylation, the process that was initiated by cyclin D/CDK4 complex. Further substrates for cyclin E/CDK2 as well as for cyclin A/CDK2 are proteins involved in DNA replication (e.g. origin-recognition complex – ORC, minichromosome maintenance proteins –MCMs, CDC6, all of which assemble into preinitiation complexes). Once cell enters S phase, cyclin A/CDK2 inactivates E2F by phosphorylation. All cyclin D-, E-, A-dependent kinases are additionally negatively regulated by a distinct "Cip/Kip" family of CDK inhibitors ($p21^{CIP1}$, $p27^{KIP1}$, and $p57^{KIP2}$). p21 is shown

to be induced by the tumor suppressor p53. p53 in turn is regulated by the murine double minute clone 2 protein (MDM2) (for details see chapter 2.3).

Studies over the last decade have indicated that most human cancer cells sustain mutations that affect the functions of RB and p53, either by disabling their genes directly or by targeting genes that act to prevent their proper function (for review see Sherr, 2000).

2.2 E2F Activity

E2F protein was first discovered as a **f**actor that interacts with the adenovirus early region 2 (**E2**) promoter (Reichel et al., 1987). E2F was shown to interact with the RB protein (Bagchi et al., 1991; Bandara et al., 1991; Chellappan et al, 1991; Chittenden et al., 1991) and this caused enormous interest in the protein. One year later it was successfully cloned (Helin et al., 1992; Kaelin et al., 1992; Shan et al., 1992). Further investigations revealed a heterodimerization partner for E2F-1 called DP-1 (Huber et al., 1993). A few more E2F family members (E2F-1 through E2F-6, as well as DP-1 and DP-2) were discovered in the following years (for review see Dyson et al., 1998 or Müller et al., 2000).

E2F-1, which differs little from other members of the family, was shown to be involved not only in cell cycle regulation, forced induction of E2F activity can override RB function and cause p53 dependent apoptosis (Qin et al, 1994). However, p53-independent apoptosis has also been observed in various cellular systems *in vitro* and *in vivo* following induction of E2F activity, but the mechanism of this phenomenon remains poorly understood. Other members of E2F family seem not to be involved in apoptosis. Additionally, E2F-1, but not other E2Fs, plays a role in cell differentiation, which is not fully explained yet (Strom et al., 1998). As already mentioned, the E2F transcription factor was identified as an important component for transcription of the adenovirus E2 gene. The DNA tumor viruses infect cell, which are quiescent, terminally differentiated that are not dividing. To create an environment appropriate for viral DNA synthesis, the viruses have to stimulate host cells to enter S phase. This stimulation is driven mostly by the viral regulatory proteins that include adenovirus E1A, SV40 T antigen, and human papillomavirus (HPV) E7. Recent developments have led to the realization that these viral proteins mediate these events through the activation of the E2F by disruption of RB/E2F complex (for review see Cress & Nevins, 1996)

2.2.1 E2F Protein Family

The E2F transcription factor family consists of at least seven distinct genes (Figure 2.2). One can divide them into two groups. E2F-1 to E2F-5 constitute the first group. DP-1 and DP-2 belong to the second one. Several forms of the DP-2 (also referred to as DP-3) protein can be produced as a result of an alternative splicing. A functional E2F transcription factor consists of a heterodimer containing an E2F and a DP polypeptide. Each of the five E2F polypeptides can heterodimerize with either DP-1 or DP-2 (DP-3). Each of these heterodimers can bind to the E2F-specific DNA sequences *in vitro* and stimulate transcription when overexpressed (for review see Johnson & Schneider, 1998). E2F-1 to E2F-3 are more closely related to each other than to E2F-4 and E2F-5.

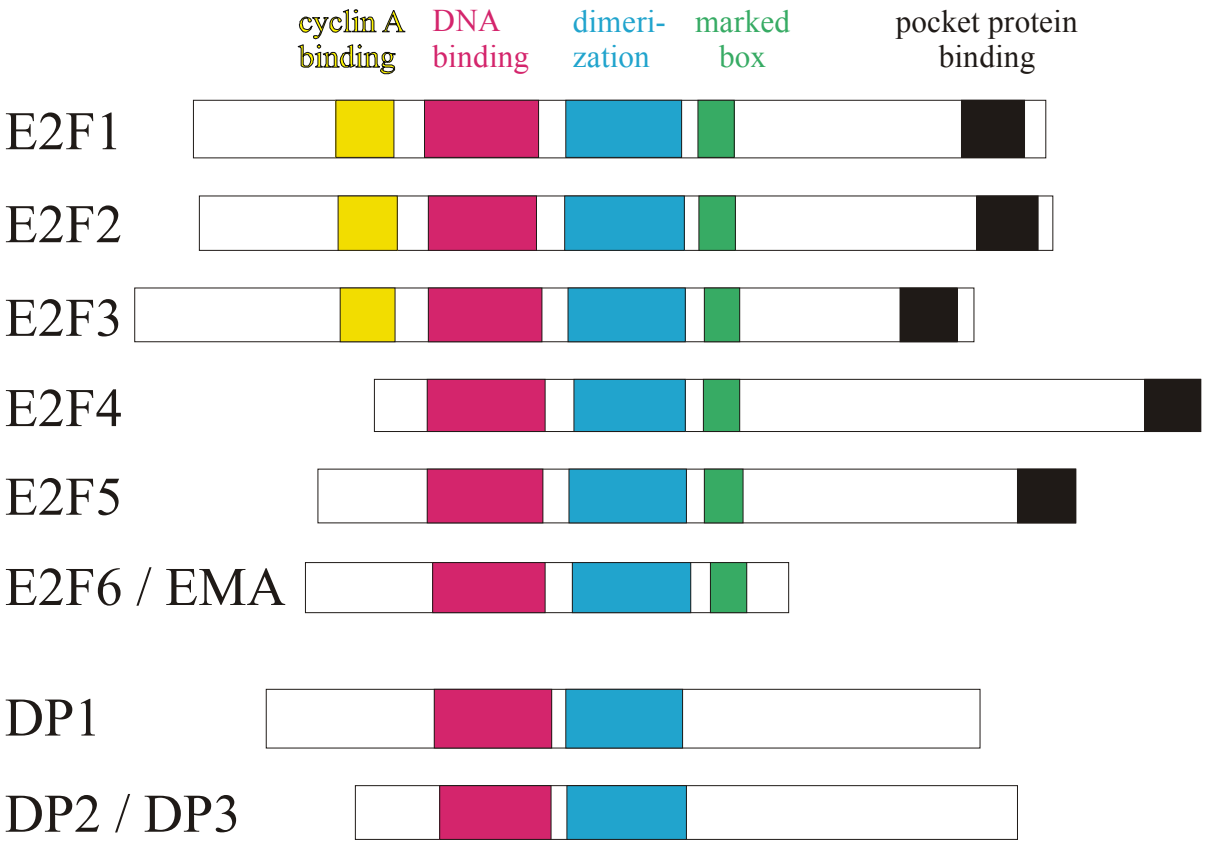


Figure 2.2. *Schematic representation of the E2F protein family.*

The domains common for E2Fs are DNA-binding domain and DP dimerization domain. Their carboxy termini contain the transcriptional activation domains as well as region involved in binding to the pocket proteins. Very well conserved is also the marked box (see next chapter). The amino termini of the first three E2Fs contain an additional region of homology not found

in E2F-4 or E2F-5, which is responsible for cyclin A binding. E2F-4 protein includes also a stretch of consecutive serine residues between the marked box and the pocket protein binding site, not found in other family members. DP polypeptides contain only some of the domains present in E2F, like DNA binding site and heterodimerization domain. Recently, an additional E2F family member was characterized and named E2F-6 or EMA (**E**2F-binding site **m**odulating **a**ctivity). E2F-6 shares homology with most of the E2F domains, but lacks the pocket protein binding domain and acidic transactivation domain. This led to an assumption that it works mainly as a repressor of E2F site-dependent transcription.

E2F polypeptides differ in their preferences for the pocket proteins (Figure 2.3). They are expressed differently during the cell proliferation, and have a tissue-specific expression. There are speculations that different E2F factors may be responsible for regulating different E2F target genes (for review see Johnson & Schneider, 1998).

Figure 2.3. *Interaction possibilities between E2Fs and pocket proteins. (Asterisk depicts not confirmed possibility, probably weak binding).*

E2F proteins were found in many different eukaryotic organisms (rat, mouse, chicken, Drosophila, wheat, tobacco) and sequence homology between them is high (Ramirez-Parra et al, 1999; Sekin et al., 1999).

2.2.2 Detailed Domain Structure of E2F-1

Initial structural analysis of the E2F-1 protein defined domains responsible for DNA binding (composed of an expected **b**asic region and **h**elix-**l**oop-**h**elix region, bHLH), transcriptional activation, and RB binding. Further analysis has also identified domains for cyclin A binding, p107 binding, homodimerization and heterodimerization (for review see Slansky & Farnham, 1996).

The E2F-1 basic region is dissimilar to that of other bHLH protein (see chapter 2.2.4 for details). Consensus sequence for the specific binding site was defined as TTTSSCGC (S=C or G). There is still an open discussion if the different members of E2F family have different specificity for particular DNA sequences (Tao et al., 1997). Bacterially expressed E2F-1 could bind to DNA without other protein partners. This binding is, however, dependent on a homodimerization domain. The heterodimer of E2F-1 and DP-1 is more active in DNA binding than are homodimeric forms of either protein (Helin et al., 1993a).

Figure 2.4. *Schematic representation of the functional domains of the human E2F-1. Gray box stands for the cyclin A binding domain (aa 67-108), red for DNA binding (120-191), blue for the homodimerization region (150-190), dark blue for heterodimerization (188-241), green for the "marked box" domain, yellow for MDM2 binding (359-437), pink for the transactivation domain (380-437), black for RB binding (409-426). Asterisks (332, 337) and hash (375) indicate phosphorylation sites. NLS stands for* ***n****uclear* ***l****ocalization* ***s****ignal (181-185).*

Analysis of point mutations of pocket proteins binding sites revealed that there is difference in the manner in which different RB homologs bind to E2Fs (Figure 2.3). RB binding domain, which is a binding domain for TBP (**T**ATA **b**inding **p**rotein) as well, partially covers transactivation domain of E2F.

E2F family members can also associate with cyclin E and cyclin A. The abundance of these complexes changes in different stages of the cell cycle. E2F/DP heterodimer is phosphorylated at distinct times in the cell cycle due to different cyclin/kinase interactions. First, in the late G_1 phase two serines on E2F-1 (amino acids 332 and 337) are phosphorylated, which leads to inhibition of RB/E2F-1 interactions (Fagan et al, 1994). The next wave of E2F phosphorylation (DP-1 becomes phosphorylated) is likely due to cyclin A/cdk2 resulting in a protein complex with reduced affinity for DNA. In G_2/M phase, E2F-1 is phosphorylated on amino acid 375 to enable high-affinity binding by RB in the subsequent G_1 phase.

E2F-1 contains domain responsible for MDM2 binding which is homologous with MDM2 binding domain of p53 (comp. chapter 2.2.4 and 2.3). However, MDM2/p53

interaction blocks p53 transactivation domain and in the case of MDM2/E2F-1 activation of the transactivation domain is observed.

There are two other E2F domains worth mentioning. First is the nuclear localization signal. Only E2F-1, E2F-2, E2F-3, and DPs contain this sequence. This is an issue of many speculations concerning regulation of E2F activity by its cellular localization. E2F-4 and E2F-5 seem to require other nuclear factors, such as their DP partners or the pocket proteins, to promote their nuclear localization.

So-called "marked box" region is well conserved in all members of E2F family. Although its cellular function is unknown, it was found to be a binding site for the adenoviral E4 protein (Helin, 1998).

2.2.3 E2F Function

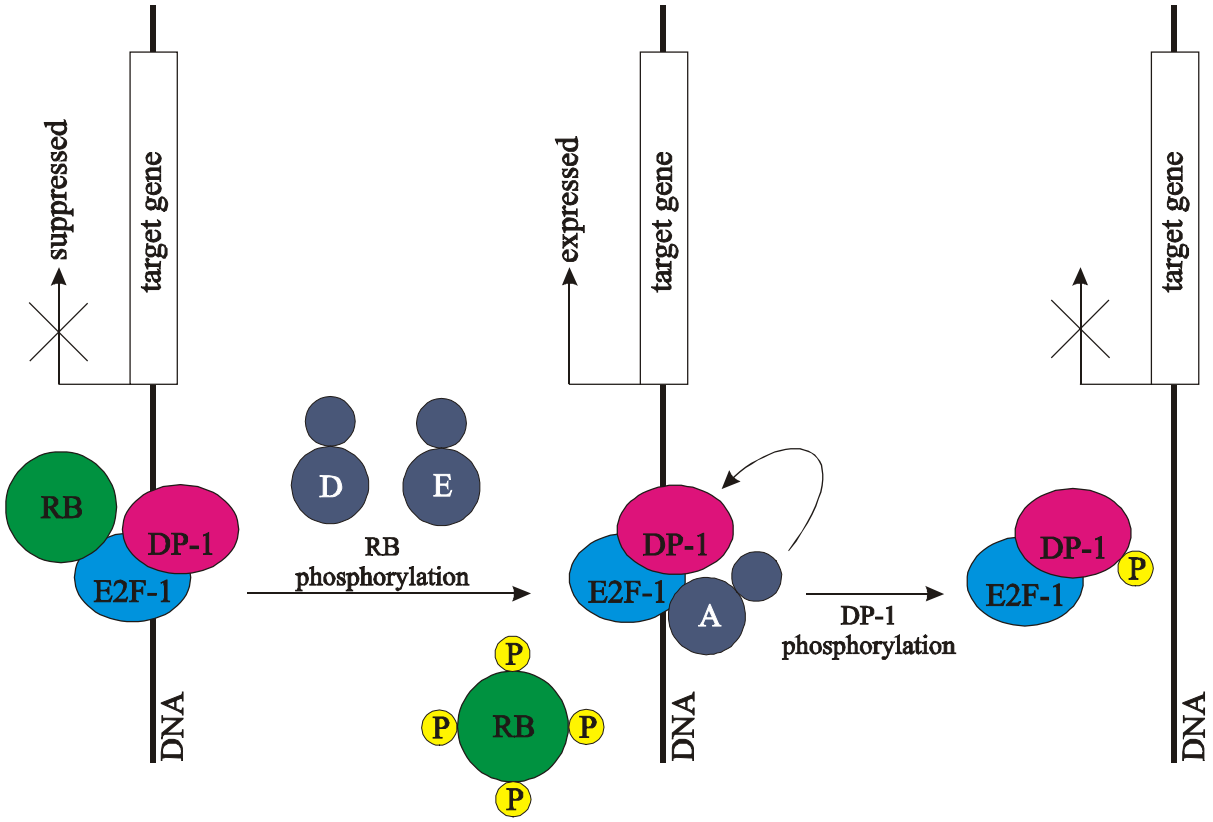


Figure 2.5. *A simplified scheme of regulation of E2F by the RB protein. When RB binds E2F at G_1 phase, expression of the target gene (for examples see Table 2.1) is suppressed. At late G_1, RB is phosphorylated by cyclin D/CDK4,6, and then by cyclin E/CDK2, which leads to the release of RB from E2F and the expression of the target gene. Subsequently, cyclin*

A/CDK2 binds E2F-1 and phosphorylates DP-1, resulting in the release of the heterodimer from DNA, and the expression of the target gene is again suppressed.

E2F activity is regulated in many different ways (E2F synthesis, E2F degradation, DNA binding by phosphorylation of DP-1, binding to pocket proteins, subcellular localization, acetylation; for review see Müller & Helin, 2000). The mechanism studied in most detail is the regulation by RB and its relatives. The pocket proteins have been reported to alter the functions of E2F by direct binding, which leads to repression of transactivation, repression of apoptosis, protection from degradation, and altered E2F-DNA binding site specificity (Figure 2.5).

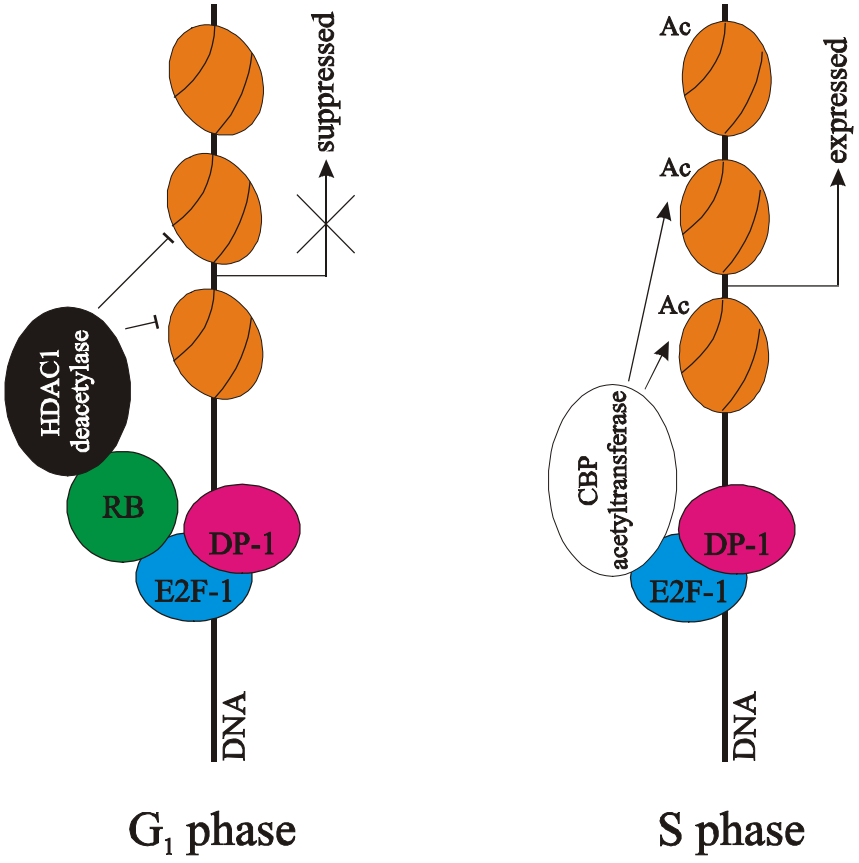


Figure 2.6. *E2F regulates transcription by interacting with acetylotransferases and deacetylases. For detail refer to text.*

Recently a more detailed explanation for the RB inhibitory role, as well as for the E2F transactivation was proposed (Brehm et al, 1999). It has been known that acetylation of lysine residues at the N-terminus of histones leads to stimulation of transcription. Neutralization of

the positive charge of these lysines is thought to "remodel" the nucleosomes along the DNA, thus allowing better access of transcription factors to the promoter. E2F was shown to be able to recruit CBP protein with **h**istone **a**cetyl**t**ransferase (HAT) activity to the promoter. Furthermore the RB protein was found to associate with a protein called HDAC1, which has an intrinsic **h**istone **d**eacetylase **ac**tivity. RB is able to recruit HDAC1 to E2F, since it can bind both proteins simultaneously. It was also shown that RB requires the HDAC1-associated deacetylase activity to repress E2F (Figure 2.6; for a review see Brehm et al, 1999).

E2F transcription factors regulate expression of a group of cellular genes that control cellular DNA synthesis and proliferation. They include cellular oncogenes, tumor suppressor genes, and other rate-limiting regulators of DNA synthesis and cell cycle progression (Table 2.1). Given the functions of these genes, it is reasonable to expect that the altered expression of the E2F factors could contribute to the development of cancer.

Table 2.1. *Partial listing of E2F target genes (Sladek, 1997).*

Functional category of target gene	***Specific genes***
DNA synthesis/nucleotide metabolism	Carbamoyl-phosphate synthase-aspartate Carbamoyltransferase-dihydroorotase
	DNA polymerase α
	Deoxycytosine kinase
	Dihydrofolate reductase (DHFR)
	Proliferating cell nuclear antigen (PCNA)
	Thymidine kinase (TK)
	Thymidylate synthetase (TS)
	Topoisomerase I
	Ribonucleotide reductase subunit M2
Cell cycle progression	*cdc2*
	Cyclin A
	Cyclin D1
	Cyclin E
Proto-oncogenes	*erb-B*
	Insulin-like growth factor I (IGF-I)
	B-*myb*
	c-*myb*
	c-*myc*
	N-*myc*
Tumor suppressor genes	RB
	p107
Others	E2F-1

2.2.4 3D Structure of E2F-4/DP-2/DNA Complex

The aim of my studies on E2F that are discussed in this PhD thesis was the preparation the E2F protein for structural investigations. During these studies an article by Zheng et al. (1999) was published, which presented the x-ray structure of a tertiary complex of E2F-4 (amino acids 11-86 fragment), DP-2 (amino acids 60-154 fragment) and a DNA (15bp fragment, containing adenovirus E2 promoter –specific binding site for E2F). The crystal structure reveals some important features of the E2F-DP-DNA interaction. Both E2F and DP have a fold related to the winged-helix DNA-binding motif, and not to the bHLH motif as previously suggested (Jordan et al., 1994). The winged-helix motif occurs in several eukaryotic transcription factors (for review see Kaufmann & Knochel, 1996). It consists of three α helices and a β sheet, each contributing to a compact hydrophobic core. DP-2 has the same overall structure as E2F-4, except that the α2 and α3 helices of DP-2 are longer by about two turns each, and E2F-4 has an amino-terminal helical extension. The E2F-4 and DP-2 structures can be well superimposed. The structure-based alignment of both proteins, however, shows that the regions have very low sequence homology. The consensus sequence of the DNA binding site is not completely symmetric. It contains a T-rich portion at one end (TTTc/gGCGCg/c). That is associated with an E2F-4 amino-terminal extension that is conserved in the E2F but not the DP family. This asymmetry in the contacts could help orient the E2F-DP heterodimer on the promoter. The complex reveals also an extensive protein-protein interface, which is mostly hydrophobic.

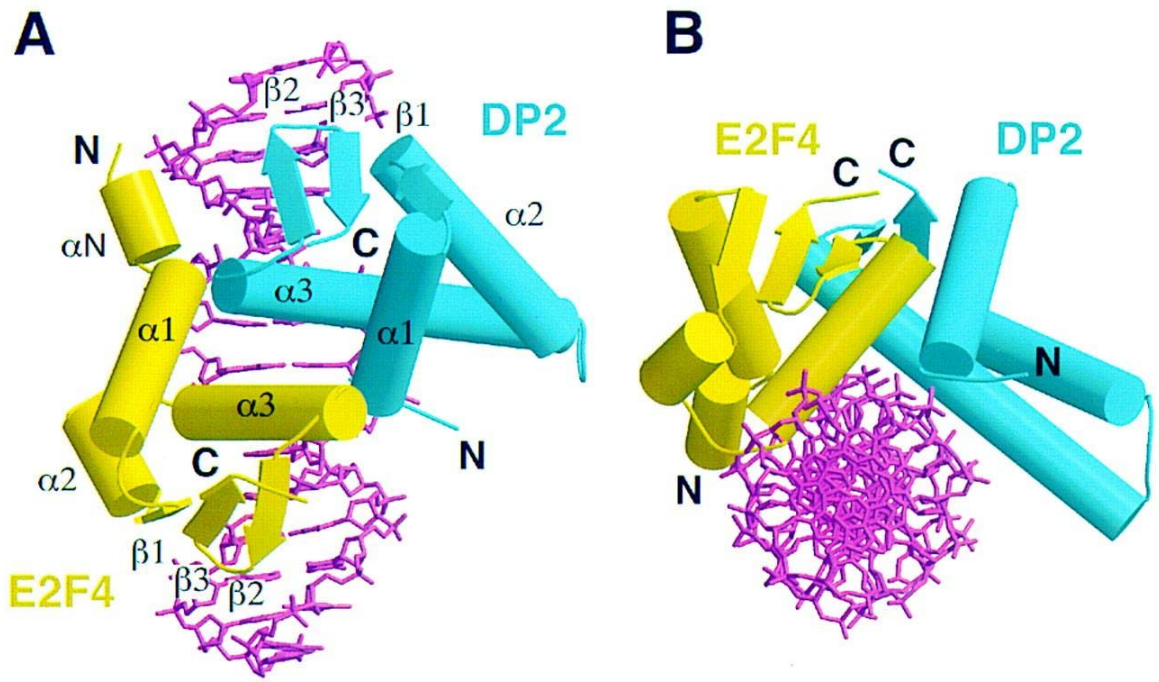


Figure 2.7. *Structure of the E2F-4/DP-2 in complex with a sequence specific DNA fragment. (A) DNA axis aligned vertically. (B) View looking down the DNA axis (Zheng et al., 1999).*

The sequence comparison of all members of E2F family allows for making additional predictions about the structure of other heterodimers in this family, as well as accounts for differences and similarities in their binding specificity (Zheng et al., 1999).

2.2.5 E2F and Cancer

It has been established that RB, as well as proteins upstream of RB (cyclin D, CDK4 and p16), are frequently mutated or deregulated in human cancer. It is now assumed that deregulation of that pathway is a prerequisite for oncogenesis. E2F, however, which constitutes a central part of this pathway, is not a frequent target of mutations in cancer. Loss of RB function and gain of E2F function seem not to have equivalent consequences. RB appears to have a broader impact on cellular homeostasis and loss of its function seems to be more advantageous for tumor development. However necessary, dysfunction of RB pathway is not sufficient for tumorigenesis. Lost of RB function often leads to the p53-dependent apoptosis. Consequently, mutations in the tumor suppressor p53 are frequently observed in conjunction with RB (Nevins, 2001).

Although the E2F activity is not directly targeted in cancer, the understanding of its function is a promising tool in cancer therapy. Attempts to block the E2F activity to prevent cancer cell proliferation, as well as overexpress E2F to eliminate cancer cells by E2F-mediated apoptosis, have been described (Fueyo et al., 1998; Liu et al., 1999). It is believe that a detailed understanding of the RB/E2F and p53 pathways will be critical for engineering the new generation of therapeutic approaches to cancer.

2.3 Oncoprotein MDM2

The mdm2 gene was first cloned as a gene amplified on double minute particles in a transformed murine cell line (**m**urine **d**ouble **m**inute gene **2**). The gene product was shown to bind the p53 tumor suppressor protein and to be amplified in many sarcomas (for review see Lozano & Montes de Oca Luna, 1998).

2.3.1 MDM2 Domain Structure

Analysis of MDM2 sequence revealed some potential functional domains of the protein (Figure 2.8), and several of them were already confirmed by functional studies. Other, however, are still waiting for confirmation. The nuclear localization sequence (NLS, amino acids 181-185) should be expected because the protein acts in the nucleus. A so-called acidic domain (aa 221-272), that contains 40% glutamic acid and aspartic acid residues, was shown to be able to activate transcription. MDM2 contains also a putative zinc finger motif (ZF, aa 305-322), which is expected to be a DNA binding domain, however there are no experimental data that could support the hypothesis. RING finger domains (RF, here aa 438-478) are known to be responsible for interactions with other proteins, DNA or RNA. In fact, MDM2 was shown to bind an RNA (for review see Lozano & Montes de Oca Luna, 1998).

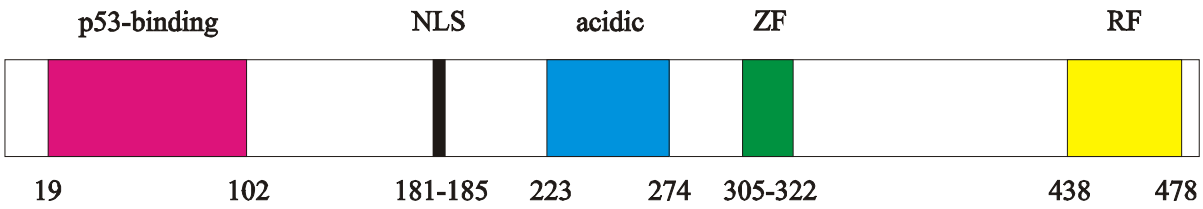


Figure 2.8. *MDM2 domains. For details refer to text.*

MDM2 is able to interact with many proteins like p53 (aa 19-102, for details see chapter 2.3.2) and RB tumor suppressors, E2F1 (aa 1-220, probably the same binding manner like p53), TATA-binding protein (TBP, aa 120-222), and the 34kDa subunit of TFIIE (aa 1-222) both *in vitro* and *in vivo*.

2.3.2 MDM2-p53 Interaction

The most important known function of the MDM2 oncoprotein is its negative regulation of the p53 function. The tumor suppressor p53 is a protein, which is maintained at low, even undetectable levels in normal cells. In response to various types of stress, such as DNA damage, the protein is stabilized and its amount in the cells rapidly increases. p53 can bind to specific DNA sequences and activate gene expression. It can lead subsequently to cell growth inhibition or apoptosis. To keep p53 level under control cells are employing several mechanisms. One of them is MDM2 expression in response to raising p53 levels, which is an

example of a simple feedback loop. Deregulation of p53 expression by mutations and other genomic alterations or by the binding of the cellular MDM2 oncoprotein was found in around 50% of human cancers (for review see Lozano & Montes de Oca Luna, 1998).

MDM2 was shown to bind p53 and inhibit its transactivation domain. It is also able to target p53 for degradation by the ubiquitin-dependent proteosome pathway (Haupt et al, 1997). A crystal structure of the 109-residue amino-terminal domain of MDM2 bound to a 15-residue transactivation domain peptide of p53 was published in 1996 (Kussie et al., 1996).

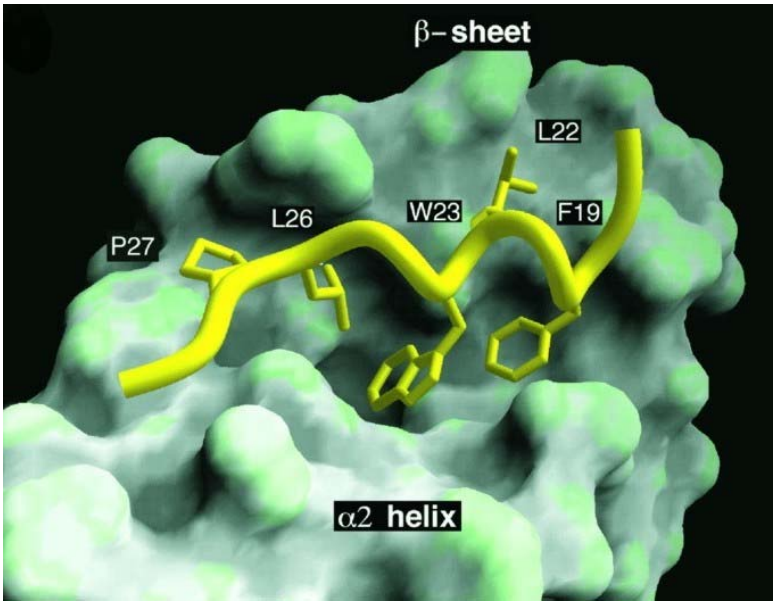


Figure 2.9. *Surface representation of the MDM2 cleft with gray concave regions highlighting its pocket-like characteristics. The p53 amino acids that interact with this surface (Phe19, Trp23, and Leu26) are shown in yellow, and are labeled (Kussie et al., 1996).*

The structure revealed that MDM2 has a deep hydrophobic cleft on which the p53 peptide binds as an amphipathic α helix. The interaction interface relies primarily on van der Waals contacts and the buried surface area is nearly all hydrophobic. Three p53 amino acids (Phe 19, Trp 23 and Leu 26) are particularly involved in binding. They insert deep into the MDM2 cleft. The same residues are involved in transactivation, which supports the hypothesis that MDM2 masks the p53 transactivation domain, thus inactivating this tumor suppressor. Recent studies have identified the E2F transcription factor and the RB protein as additional targets of the protein binding activity of MDM2, and have suggested a broader role for MDM2 in modulating cell growth controls. The N-terminal domain of MDM2 is also the binding site for the E2F-1 transcription factor. E2F-1 contains a region homologous to the

MDM2 binding portion of p53, having a similar pattern of hydrophobic amino acids (D*F*SG*LL*PEE of E2F-1 compared to T*F*SD*LW*KLL of p53, with the homologous hydrophobic residues in italics). These observations suggest that the E2F-1/MDM2 and p53/MDM2 interfaces may have common structural elements. Nevertheless, there are likely to be differences in the two complexes because Leu 26 of p53, which makes key contacts to hydrophobic MDM2 residues, is replaced by a glutamic acid in E2F1 (Kussie et al., 1996). The crystal structure thus provides a framework for the discovery of compounds that may prevent the inactivation of the p53 tumor suppressor by the MDM2 oncogene in cancer.

2.4 IGF System

The **i**nsulin-like **g**rowth **f**actors (IGF-I and IGF-II) are potent mitogens. IGF-I, a 70-amino acid-protein structurally similar to insulin (ca. 50% homology), promotes cell proliferation and differentiation in multiple tissues. Most of its effects are mediated by the Type **I IGF r**eceptor (IGF-IR), a heterotetramer that has tyrosine kinase activity and phosphorylates **i**nsulin **r**eceptor **s**ubstrates (IRS-1 and 2) which leads to the activation of two downstream signaling cascades: the **m**itogen-**a**ctivated **p**rotein kinase (MAP) and the **p**hosphatidylinositol **3**-**k**inase (P3K) cascades. The primary regulator of IGF-I expression is **g**rowth **h**ormone (GH), however the developmental expression of IGF-I in various tissues precedes that of GH, supporting an independent role of IGF-I in an embryonic and fetal life. The level of free IGF in a system is modulated by the degree of binding to **IGF b**inding **p**roteins (IGFBPs, see chapter 2.4.1). Most of the IGF molecules in serum are found in a 150 kDa ternary complex formed by an IGF, IGFBP-3 or 5 and a glycoprotein known as the **a**cid **l**abile **s**ubunit (ALS). Less than 1% of IGFs circulate in the free forms. The IGF/IGFBP complex is acted upon by proteases at the target organ, whereby IGF is released and is available for biological actions. The GH/IGF-I axis is the primarily regulator of postnatal growth while IGF-II, which is relatively independent of GH, appears to have an important role during fetal development (Khandwala et al., 2000).

2.4.1 IGFBP Superfamily

IGFBPs constitute a family of proteins whose major function is binding and regulation of the IGF hormones. The family comprises six proteins (IGFBP-1 to –6) that bind to IGFs with high affinity and a group of **IGFBP-r**elated **p**roteins (IGFBP-rP 1-9), which bind IGFs

with lower affinity. The proteins are ubiquitously produced in all tissues, with each tissue having specific levels of certain IGFBPs. A key conserved structural feature among six IGFBPs is the high number of cysteines (16-20 cysteines), clustered at the N-terminus (12 cysteines) and the C-terminus. The proteins share a high degree of similarity (36%) in their primary protein structure, particularly in their N- and C-terminal regions. It has been therefore postulated that these regions participate in the high-affinity binding of IGFs. Consistent with this hypothesis, IGFBP-rPs, which are homologous to the IGFBPs only in the N-terminus, were shown to have at least 20-25-fold lower affinity for IGFs (for review see Hwa et al., 1999). Most IGFBPs show higher affinity for IGF-I than IGF-II, except IGFBP-6, which has 100-fold higher affinity for IGF-II versus IGF-I.

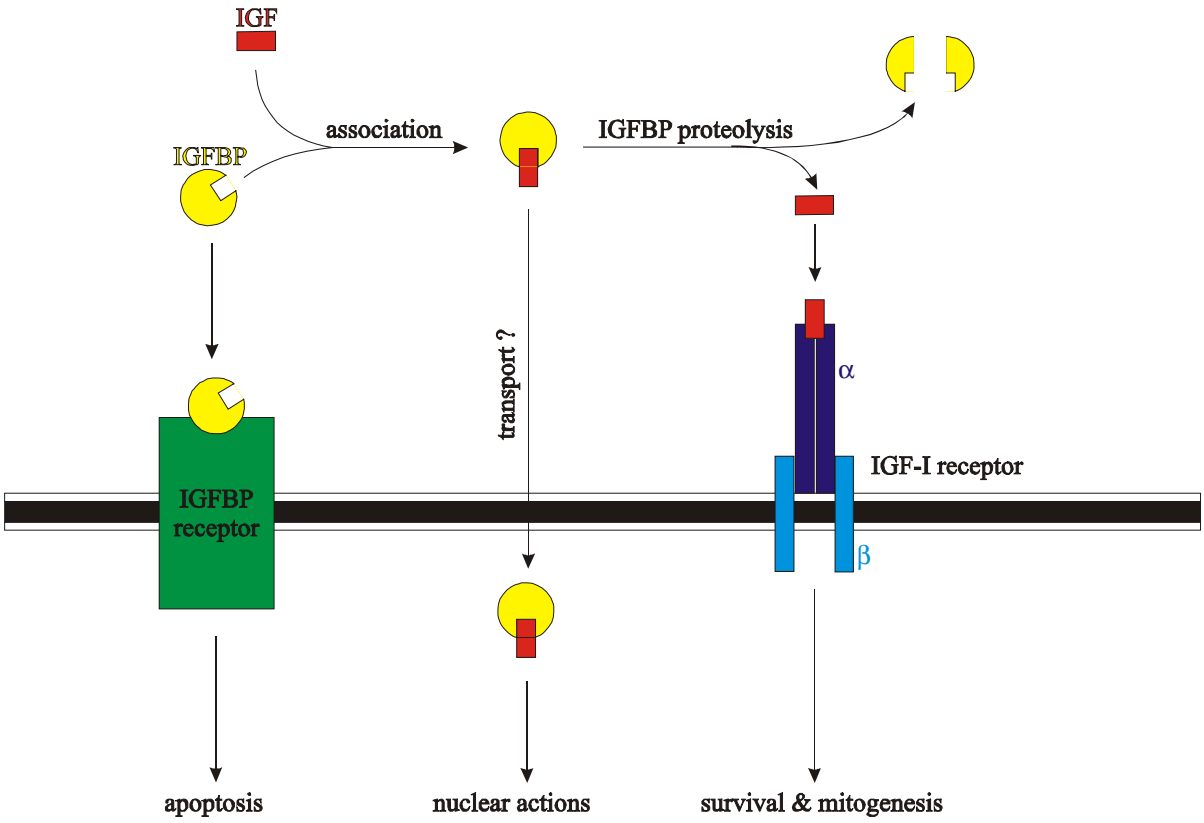


Figure 2.10. *The IGF-IGFBP axis. IGFBP binds IGF with high affinity, regulating the bioavailability of free IGF. Specific IGFBP proteases cleave IGFBP and thereby regulate levels of free IGF as well as IGFBP. IGFBPs are involved in both IGF-dependent and IGF-independent actions.*

The high-affinity IGFBPs modulate IGF bioavailability by undergoing proteolysis and generating fragments with reduced or no affinity for IGFs. Categories of IGFBP proteases

include kallikreins, cathepsins, and matrix metalloproteinases (Wetterau et al., 1999). IGFBPs not only regulate IGF action but also appear to mediate IGF-independent actions, including inhibition or enhancement of cell growth and induction of apoptosis. Recently, the presence of specific cell-surface IGFBP receptors were discovered. IGFBP-3 and –5 have recently been shown to be translocated into the nucleus compatible with having nuclear localization sequence (NLS) in their mid-region. This raises the possibility that nuclear IGFBP may directly control gene expression. IGFBPs were also shown to bind to important viral oncoproteins like HPV oncoprotein E7. This implies additional roles for IGFBPs in the pathways of cell proliferation, apoptosis, and malignant transformation.

Full length IGFBP-5 is a 29kDa protein. It is expressed mainly in the kidney, and is found in substantial amounts in connective tissues. Unlike other IGFBPs, IGFBP-5 strongly binds to bone cells because of its high affinity for hydroxyapatite. *In vitro* studies showed that IGFBP-5 stimulated IGF-I actions when compared to IGF-I alone. This stimulation is particularly evident in bone cells. IGFBP-5 is believed to be bound to the cell membrane or to extracellular matrix to cause this potentiating effect. The binding of IGF-I to the matrix-bound IGFBP-5 facilitates subsequent binding of IGF-I to its receptors. The regulatory mechanisms for IGFBP-5 are still under investigation.

The IGFs, with their potent mitogenic and antiapoptotic effects, have been widely studied for their role in cancer (for review see Khandwala et al., 2000). Serum IGF-I and IGFBP-3 have been even proposed as candidate markers for early detection of some cancers. Recombinant forms of IGFBPs as well as compounds, which could disrupt their interaction with IGFs are thus potential therapeutic tools.

3 Methods for Structural Studies

In this chapter two most powerful techniques for structural studies will be shortly presented. They are, to date, the only methods, which can solve protein structure in three dimensions. NMR spectroscopy, which has the disadvantage being very time consuming and restricted for small molecular weight proteins (up to 30kDa) in protein structure determination, has, however, many advantages in comparison to x-ray crystallography, as it can be useful for dynamics studies and can provide us with many other useful information about the protein in the solution.

3.1 X-ray Crystallography

The first prerequisite for solving the three-dimensional structure of a protein by x-ray crystallography is a well-ordered crystal that will diffract x-rays strongly. The crystallographic method depends upon directing a beam of x-rays onto a regular, repeating array of many identical molecules so that the x-rays are diffracted from it in a pattern, a **diffraction pattern**, from which the structure of an individual molecule can be retrieved. The repeating unit forming the crystal is called the **unit cell**. Each unit cell may contain one or more molecules.

There are several techniques for setting up crystallization experiments including sitting drop vapor diffusion, hanging drop vapor diffusion, sandwich drop, batch, microbatch, under oil, microdialysis, and free interface diffusion. Sitting and hanging drop methodologies are very popular because they are easy to perform, require a small amount of sample, and allow only a large amount of flexibility during screening and optimization.

Using the **sitting drop technique** (Figure 3.1.A) one places a small (1 to 40 microliters) droplet of the sample mixed with crystallization reagent on a platform in vapor equilibration with the reagent. The initial reagent concentration in the droplet is less than that in the reservoir. Over time the reservoir will pull water from the droplet in a vapor phase such that an equilibrium will exist between the drop and the reservoir. During this equilibration process the sample is also concentrated, increasing the relative supersaturation of the sample in the drop. The advantages of the sitting drop technique include speed and simplicity. The disadvantages are that crystals can sometimes adhere to the sitting drop surface making removal difficult. This disadvantage can turn into an advantage where occasionally the surface of the sitting drop can assist in nucleation. The sitting drop is an excellent method for

screening and optimization. During production, if sticking is a problem, sitting drops can be performed in the sandwich box set up.

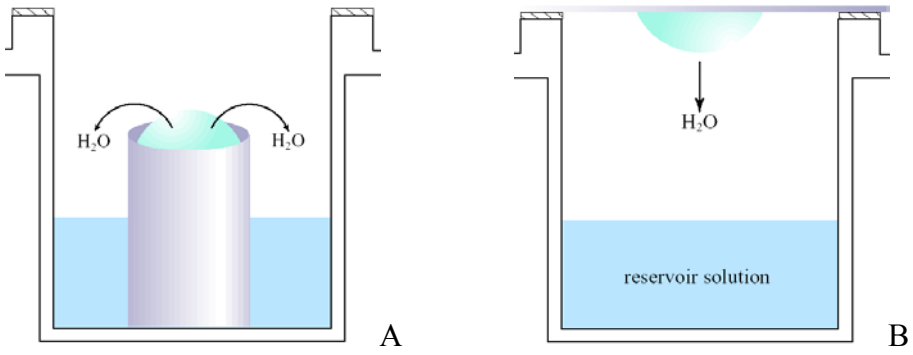


Figure 3.1. (A) *Sitting drop and* (B) *hanging drop crystallization techniques.*

Using the **hanging drop technique** (Figure 3.1.B) one places a small (1 to 40 microliters) droplet of the sample mixed with crystallization reagent on a siliconized glass cover slide inverted over the reservoir in vapor equilibration with the reagent. The initial reagent concentration in the droplet is less than that in the reservoir. Over time the reservoir will pull water from the droplet in a vapor phase such that an equilibrium will exist between the drop and the reservoir. During this equilibration process the sample is also concentrated, increasing the relative supersaturation of the sample in the drop. The advantages of the hanging drop technique include the ability to view the drop through glass without the optical interference from plastic, flexibility, reduced chance of crystals sticking to the hardware, and easy access to the drop. The disadvantage is that a little extra time is required for set ups.

When the primary beam from an x-ray source strikes the crystal, most of the x-rays travel straight through it. Some, however, interact with the electrons on each atom and cause them to oscillate. The oscillating electrons serve as a new source of x-rays, which are emitted in almost all directions. We refer to this rather loosely as **scattering**. When atoms and hence their electrons are arranged in a regular three-dimensional array, as in crystal, the x-rays emitted from the oscillating electrons interfere with one another. In most cases, these x-rays, colliding from different directions, cancel each other out; those from certain directions, however, will add together to produce diffracted beams of radiation that can be recorded as a pattern on a detector. Diffraction by a crystal can be regarded as the reflection of the primary beam by sets of parallel planes. The relationship between the reflection angle, θ, the distance between the planes, d, and the wavelength, λ, is given by **Bragg's law**:

$$2d \sin\theta = \lambda$$

This relation can be used to determine the size of the unit cell. Each atom in a crystal scatters x-ray in all directions, and only those that positively interfere with one another, according to Bragg's law, give rise to diffracted beams that can be recorded as a distinct **diffraction spot** above background. Each diffraction spot is the result of interference of all x-rays with the same diffraction angle emerging from all atoms. For a typical protein crystal, myoglobin, each of the about 20000 diffracted beams that have been measured contains scattered x-rays from each of the around 1500 atoms in the molecule. To extract information about individual atoms from such a system the **Fourier transform** is employed. Each diffracted beam, which is recorded as a spot on the detector, is defined by three properties: the amplitude, which can be measured from intensity of the spot; the wavelength, which is set by the x-ray source; and the phase, which is lost in x-ray experiments. All three properties for all of the diffracted beams have to be known to determine the position of the atoms. This is so-called **phase problem** in x-ray crystallography. There are two methods, to date, which can help to circumvent the phase problem.

Multiple isomorphous replacement (MIR), requires the introduction of new x-ray scatterers into the unit cell of the crystal. These additions should be, for example, heavy atoms, so that they can make a significant contribution to the diffraction pattern. Following heavy-metal substitution, some spots measurably increase in intensity, others decrease, and many show no detectable difference. The intensity differences can be used to deduce the positions of the heavy-atoms in the crystal unit cell. Fourier summations of these intensity differences give maps of the vectors between the heavy atoms, the so-called **Patterson maps**. From these vector maps, the atomic arrangement of the heavy atoms can be deduced. From the positions of the heavy metals in the unit cell, one can calculate the amplitudes and phases of their contribution to the diffracted beams of the protein crystals containing heavy metals. Subsequently the phase of the protein can be calculated.

Phase information can also be obtained by **multiwavelength anomalous diffraction (MAD)** experiments. For certain x-ray wavelengths, the interaction between the x-rays and the electrons of an atom causes the electrons to absorb the energy of the x-ray. This causes a change in the x-ray scattering of the atom, called **anomalous scattering**. The intensity differences obtained in the diffraction pattern by illuminating such a crystal by x-rays of different wavelengths can be used in a way similar to the method of MIR to obtain the phases of the diffracted beams. The MAD method requires access to synchrotron radiation since the different wavelengths are used.

The amplitudes and the phases of the diffraction data from the protein crystals are used to calculate an **electron-density map** of the repeating unit of the crystal. This map then has to be interpreted as a polypeptide chain with a particular amino acid sequence. The quality of the map depends on the **resolution** of the diffraction data, which in turn depends on how well-ordered the crystals are. The initial model will contain errors, which can be subsequently removed by crystallographic **refinement** of the model. In this process the model is changed to minimize the difference between the experimentally observed diffraction amplitudes and those calculated for a hypothetical crystal containing the model instead of the real molecule. This difference is expressed as an **R factor**, residual disagreement, which is 0.0 for exact agreement and around 0.59 for total disagreement. In general R factor is between 0.15 and 0.20 for a well-determined protein structure. In refined structures at high resolution (around 2Å) there are usually no major errors in the orientation of individual residues. Hydrogen bonds within the protein and to bound ligands can be identified with a high degree of confidence.

The introduction presented here was adapted from Branden & Tooze (1999). For more details about crystallization procedures refer to McPherson (1999), and for the basis of x-ray crystallography refer to Drenth (1994) or Giacovazzo (1992).

3.2 NMR Spectroscopy

When protein molecules are placed in a strong magnetic field, the spin of their hydrogen atoms aligns along the field. This equilibrium alignment can be changed to an excited state by applying radio frequency (RF) pulses to the sample. When the nuclei of the protein molecule revert to their equilibrium state, they emit RF radiation that can be measured. The exact frequency of the emitted radiation from each nucleus depends on the molecular environment of the nucleus and is different for each atom, unless they are chemically equivalent and have the same molecular environment. These different frequencies are obtained relative to a reference signal and are called **chemical shifts**. The nature, duration, and combination of applied RF pulses can be varied enormously, and different molecular properties of the sample can be probed by selecting the appropriate combination of pulses.

One-dimensional spectra of protein molecules contain overlapping signals from many hydrogen atoms because the differences in chemical shifts are often smaller than the resolving power of the experiment. This problem has been bypassed by designing experimental conditions that yield a two-dimensional NMR spectrum. The diagonal in such a diagram

corresponds to a normal one-dimensional NMR spectrum. The peaks off the diagonal result from interactions between hydrogen atoms that are close to each other in space. By varying the nature of the applied RF pulses these off-diagonal peaks can reveal different types of interactions. A **COSY** (**co**rrelation **s**pectroscop**y**) experiment gives peaks between hydrogen atoms that are covalently connected through one or two other atoms. An **NOE** (**n**uclear **O**verhauser **e**ffect) spectrum, on the other hand, gives peaks between pairs of hydrogen atoms that are close together in space even if they are from amino acid residues that are quite distant in the primary sequence.

Two-dimensional NOE spectra, by specifying which groups are close together in space, contain three-dimensional information about the protein molecule. It is far from trivial, however, to assign the observed peaks in the spectra to hydrogen atoms in specific residues along the polypeptide chain because the order of peaks along the diagonal has no simple relation to the order of amino acids along the polypeptide chain. This problem has been solved by **sequential assignment**, which is based on the differences in the number of hydrogen atoms and their covalent connectivity in the different amino acid residues. Each type of amino acid has a specific set of covalently connected hydrogen atoms that will give a specific combination of **cross-peaks**, a "fingerprint", in a COSY spectrum. From the COSY spectrum it is therefore possible to identify the H atoms that belong to each amino acid residue and, in addition, determine the nature of the side chain of that residue. However, the order of these fingerprints along the diagonal has no relation to the amino acid sequence of the protein. The sequence-specific assignment, however, can be made from NOE spectra that record signals from H atoms that are close together in space. These signals in the NOE spectra make it possible to determine which fingerprint in the COSY spectrum comes from residue adjacent to the one previously identified.

The final result of the sequence-specific assignment of NMR signals is a list of **distance constraints** from specific hydrogen atoms in one residue to hydrogen atoms in a second residue. This list immediately identifies the secondary structure elements of the protein molecule because both α helices and β sheets have very specific sets of interactions of less than 5Å between their hydrogen atoms. It is also possible to derive models of the three-dimensional structure of the protein molecule. However, usually a set of possible structures rather that a unique structure is obtained.

The introduction to NMR spectroscopy presented above was adapted from Branden & Tooze (1999).

3.2.1 SAR by NMR

The "SAR (structure-activity relationships) by NMR" is a method in which small organic molecules that bind to proximal subsites of a protein are identified, optimized, and linked together to produce high-affinity ligands.

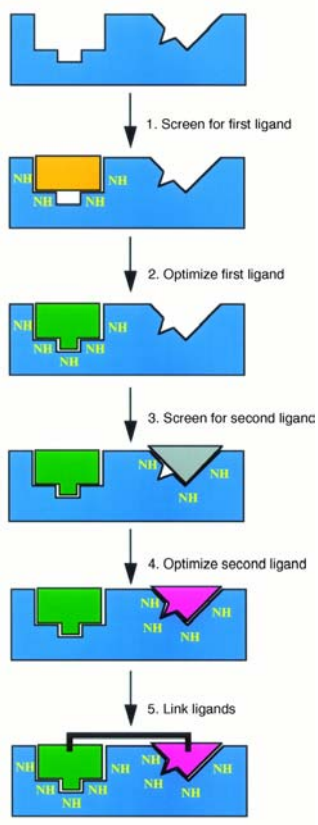


Figure 3.2. *The SAR by NMR method (from Shuker et al., 1996).*

The first step of the method is screening a library of low molecular weight compounds to identify molecules that bind to the protein. Binding is determined by the observation of ^{15}N- or ^{1}H-amide chemical shift changes in two-dimensional ^{15}N-heteronuclear single-quantum correlation (^{15}N-HSQC) spectra upon the addition of a ligand to a uniformly ^{15}N-labeled protein. Once a lead molecule is identified, analogs are screened to optimize binding to this site. Subsequently the search for another ligand that interacts with a nearby site is performed. From an analysis of the chemical shift changes, the approximate location of the second ligand can be defined. Optimization of the second ligand is then carried out by screening structurally related compounds. When two "lead" fragments have been selected, their location and orientation in the ternary complex are determined experimentally either by NMR spectroscopy or by x-ray crystallography. Finally, on the basis of this structural information, compounds are synthesized in which the two fragments are linked together with the goal of producing a high affinity ligand (Shuker et al., 1996).

The advantage of the use of ^{15}N-HSQC spectra is ability to detect the binding of small, weakly bound ligands to ^{15}N-labeled target protein. Because of the ^{15}N spectral editing, no signal from the ligand is observed. Another advantage is the ability to rapidly determine the different binding site locations of the fragments, which is critical for interpreting structure-activity relationships and for guiding the synthesis of linked compounds. However, SAR by NMR method is limited by the solubility of compounds at milimolar concentrations and is applicable only to small biomolecules (MW < 30kDa) that can be obtained in large quantities, detectable for NMR.

4 Materials and Laboratory Methods

4.1 Materials

All chemicals used in the work were supplied from Merck (Darmstadt, FRG) or Sigma (Deisenhofen, FRG) unless otherwise indicated.

4.1.1 Organisms

bacterial strains: *Escherichia coli* DH5α, BL21, BL21(DE3)

insect cell line for baculovirus system: *Spodoptera frugiperda* Sf9

4.1.2 Plasmids and Viruses for Protein Expression

E2F:

Plasmids for protein overexpression in *E.coli* were a kindly gift from Dr. Kristian Helin (Milan, Italy):

#	construct	description	reference
1	pGST2TDP1	full length DP-1, fused to GST	Helin et al., 1993a
2	pGST20TE2F1	full length E2F-1, fused to GST	Helin et al., 1993b
4	pREP4, His6E2F1	His-tagged full length E2F-1 (QIAGEN expression system)	not published
5	pT5TE2F1	tubulin-tagged full length E2F-1	Huber et al., 1993
6	pGST-E2F1 (P17)	E2F-1 (aa: 90-437) fragment, fused to GST	Helin et al., 1992
9	pGST-E2F1 (P19N)	E2F-1 (aa: 90-191) fragment, fused to GST	Helin et al., 1992
10	pGST-E2F1 (P20N)	E2F-1 (aa: 90-238) fragment, fused to GST	Helin et al., 1992

Baculoviruses for protein overexpression in insect cells were kindly provided by Prof. Jonathan M. Horowitz (Durham, North Carolina, USA):

#	description	reference
14	hemagglutinin(HA)-tagged full length E2F-1	Tao et al., 1997
15	hemagglutinin(HA)-tagged full length E2F-4	Tao et al., 1997
16	hemagglutinin(HA)-tagged full length DP-1	Tao et al., 1997
17	not tagged full length DP-2	Tao et al., 1997

MDM2:

pQE40 (QIAGEN, FRG) with inserted MDM2 fragment (1-118 aa).
(Stoll et al., 2001)

IGFBP-5:

pQE30 (QIAGEN, FRG) with inserted IGFBP-5 fragment (40-92 aa).

(Kalus et al., 1998)

4.1.3 Enzymes, Antibodies and Other Proteins

antibodies:

against	name	catalog no.	supplier
E2F-1	KH95	sc-251	Santa Cruz Biotechnology, Inc., USA
E2F-4	D-3	sc-6851	Santa Cruz Biotechnology, Inc., USA
DP-1	TFD10.2	66201A	Becton Dickinson, FRG
DP-2	G-12	sc-6849	Santa Cruz Biotechnology, Inc., USA
hemagglutinin, HA.11	16B12	MMS-101P	BAbCo, Richmond, CA, USA
goat anti-mouse IgG	-	sc-2005	Santa Cruz Biotechnology, Inc., USA

Molecular Weight Marker for SDS-PAGE Electrophoresis (NEB, FRG):

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
β-lactoglobulin A	Bovine milk	25 000
lysozyme	Chicken egg white	16 500
aprotinin	Bovine lung	6 500

other proteins:

- hen egg white lysozyme
- RNaseA
- DNaseI
- factor Xa (NEB, FRG)
- thrombin (Sigma, FRG)

4.1.4 Other Chemicals

Antibiotika:

- Ampicillin
- Chloramphenicol
- Kanamycin

Protease Inhibitors:

- Complete Protease Inhibitors Cocktail (Roche, FRG)

Isotopically Enriched Chemicals:

- Deuterium oxide, D_2O 99%, 99.99% (Campro Scientific, Berlin, FRG)
- ^{15}N-Ammonium chloride, NH_4Cl 99.9% (Campro Scientific, Berlin, FRG)

Other Chemicals:

- Acetic acid
- Acrylamide
- L-Arginine
- Ammonium chloride, NH_4Cl
- Ammonium persulfate, APS
- Biotin
- Boric acid, H_3BO_3
- Calcium chloride, $CaCl_2$
- Citric acid
- Cobalt (II) chloride, $CoCl_2$
- Coomassie Brillant Blue R-250
- Copper (II) chloride, $CuCl_2$
- Dimethylsulfoxide, DMSO
- Dipotassium hydrogenphosphate, K_2HPO_4
- Disodium hydrogenphosphate, Na_2HPO_4
- 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, $MgCl_2$
- Magnesium sulfate, $MgSO_4$
- Manganese (II) chloride, $MnCl_2$
- β-Mercaptoethanol, β-ME
- Methanol
- N,N'-Methylenbisacrylamide
- Nonidet P-40, NP-40
- Potassium chloride, KCl
- Potassium dihydrogenphosphate, KH_2PO_4
- Silver nitrate, $AgNO_3$
- Sodium acetate
- Sodium azide, NaN_3
- Sodium carbonate, Na_2CO_3
- Sodium chloride, NaCl
- Sodium dihydrogenphosphate, NaH_2PO_4
- Sodium dodecylsulphate, SDS
- Sodium hydrogencarbonate, $NaHCO_3$
- Sodium hydroxide, NaOH
- Sodium molybdate, Na_2MoO_4
- Sodium thiosulfate

- N,N,N',N'-Tetramethylenethylendiamine, TEMED
- Thiamin
- Tricine
- Trifluoroethanol, TFE
- Tris-(hydroxymethyl)-aminomethane, TRIS
- Triton X-100
- Tryptone
- Yeast Extract
- Zinc acetate, $Zn(Ac)_2$

4.1.5 Buffers and Media

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

LB Medium:

Tryptone	10g/l
Yeast Extract	5g/l
NaCl	5g/l

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

Minimal Medium (MM) for Uniform Enrichment with ^{15}N:

Stock solutions:

1. thiamin, 1%
2. antibiotika
3. $MgSO_4$, 1M
4. Zn-EDTA solution:

EDTA	5mg/ml
$Zn(Ac)_2$	8.4mg/ml

   Dissolved separately in small water volumes, then mixed together.
5. trace elements solution:

H_3BO_3	2.5g/l
$CoCl_2*H_2O$	2.0g/l

$CuCl_2*H_2O$	1.13g/l
$MnCl_2*2H_2O$	9.8g/l
$Na_2MoO_4*2H_2O$	2.0g/l

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

6. glucose, 5g/25ml, separately autoclaved.

For 1liter medium:

1. mixture was prepared:

stirring element	
NaCl	0.5g
trace elements solution	1.3ml
citric acid monohydrate	1g
ferrous citrate	36mg (dissolved in 120µl conc. HCl, heated)
KH_2PO_4	4.02g
$K_2HPO_4*3H_2O$	7.82g
Zn-EDTA solution	1ml
NH_4Cl or $^{15}NH_4Cl$	1g

2. pH was adjusted to 7.0 with NaOH
3. the mixture was autoclaved
4. 25ml separately autoclaved glucose was added
5. other compounds were added (previously sterile filtered):

thiamin	560µl
antibiotika	(half of the usual amount for LB-medium)
$MgSO_4$, 1M	2ml

Medium for Sf9 cells:

Sf-900 II SFM (Gibco, UK) supplemented with fetal bovine serum (FBS).

IPTG stock solution:

IPTG was dissolved in water (2.38g/10ml) to the endconcentration of 1M. The stock solution was sterile filtered and stored in aliquots at –20°C until used. The stock solution was diluted 1:1000 when added to the medium, unless otherwise indicated.

Ampicillin stock solution:

Ampicillin was dissolved in water (1g/10ml) to the endconcentration of 100mg/ml. The stock solution was sterile filtered and stored in aliquots at –20°C until used. The stock solution was diluted 1:1000 when added to the medium.

Kanamycin stock solution:

Kanamycin was dissolved in water (0.5g/10ml) to the endconcentration of 50mg/ml. The stock solution was sterile filtered and stored in aliquots at –20°C until used. The stock solution was diluted 1:1000 when added to the medium.

Chloramphenicol stock solution:

Chloramphenicol was dissolved in ethanol (0.5g/10ml) to the endconcentration of 50mg/ml. The stock solution was sterile filtered and stored in aliquots at –20°C until used. The stock solution was diluted 1:1000 when added to the medium.

Crystallization Buffer:

5mM tris-HCl, pH 8.0	2.42g/4l
50mM NaCl	11.68g

Phosphate-Buffered Saline (PBS) Buffer:

10mM Na_2HPO_4*$2H_2O$, pH 7.3	1.78g/l
1.8mM KH_2PO_4	1.36g
140mM NaCl	8.18g
2.7mM KCl	0.2g
0.05% NaN_3	0.5g

4.1.6 Laboratory Equipment

Consumables:

Centripreps YM3, YM10	Amicon, Witten, FRG
Dialysis tubing Spectra/Por MW 3500, 10000	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

Chromatography equipment, columns and media:

ÄKTA explorer 10	Amersham Pharmacia, Freiburg, FRG
Peristaltic pump P-1	Amersham Pharmacia, Freiburg, FRG
Fraction collector RediFrac	Amersham Pharmacia, Freiburg, FRG
Recorder REC-1	Amersham Pharmacia, Freiburg, FRG
UV flow through detector UV-1	Amersham Pharmacia, Freiburg, FRG
BioloLogic LP System	Biorad, FRG
HiLoad 16/60 Superdex S30pg, S200pg	Amersham Pharmacia, Freiburg, FRG
HiLoad 26/60 Superdex S75pg	Amersham Pharmacia, Freiburg, FRG
HiLoad 10/30 Superdex S75pg	Amersham Pharmacia, Freiburg, FRG
Mono Q HR 5/5, 10/10	Amersham Pharmacia, Freiburg, FRG
Mono S HR 5/5, 10/10	Amersham Pharmacia, Freiburg, FRG
NiNTA-agarose	QIAGEN, FRG
Buthyl Sepharose 4 FF	Amersham Pharmacia, Freiburg, FRG
Q-Sepharose FF	Amersham Pharmacia, Freiburg, FRG
SP-Sepharose FF	Amersham Pharmacia, Freiburg, FRG
Glutathione Sepharose	Amersham Pharmacia, Freiburg, FRG

Miscellaneous:

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	MPI für Biochemie, 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, Freiburg, FRG
Ultrafiltration cells, 10ml, 50ml, 200ml	Amicon, Witten, FRG
Vortex	Cenco, FRG
X-ray generator RU2000, 45kV, 120mA	Rigaku, Tokyo, Japan

4.2 Molecular Biology Techniques

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

4.2.1 Electroporation

Protocol for Electrocompetent Cells:

1. Bacteria were streaked on an LB agar plate, and incubated at 37°C overnight.
2. 50ml of LB medium in a 250ml flask were inoculated with a single colony from the LB plate and incubated at 37°C with shaking (200rpm) overnight.
3. 1l of LB medium in a 3l flask was inoculated with the 50ml overnight culture. The culture was grown in shaking (200rpm) incubator at 37°C until the OD_{600} 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 2000G for 15min. at 0 – 4°C.
5. Supernatant was decanted, and bottles put back on ice. The cell pellet in each bottle was resuspended in approximately 500ml of cold (0 – 4°C) sterile water, and subsequently centrifuged like before.

6. The cells in each bottle were washed again with 250ml of cold sterile water, and centrifuged.
7. The cell pellet in each bottle was then resuspended in 20ml cold sterile 10% glycerol and transferred to a chilled, sterile, 50ml centrifuge tube. Centrifugation followed at 4000G for 15min. at 0 – 4°C.
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 (40µl) to pre-chilled 1.5ml tubes and frozen immediately in liquid nitrogen. The aliquots were kept at – 80°C ready for use.

Transformation of the Electrocompetent Bacteria:

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

4.2.2 Bacterial Cultures

Bacterial Culture in LB medium:

1. 50ml LB were inoculated with a fresh single bacterial colony and incubated overnight at 37°C with vigorous shaking (280rpm) in 100ml flask.
2. Pre-warmed 1l LB medium in 3l flask was inoculated with 10ml of the overnight culture, supplemented with appropriate antibiotika, and incubated at 37°C with shaking (150rpm) until the OD_{600} reached the 0.7 value.
3. Induction by IPTG addition followed. 0.1-1mM IPTG concentration was usually used. The cells were then grown until the expected OD was reached.

Bacterial Culture in MM:

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

4.3 Tools of Biochemistry

All biochemical methods that are not mentioned here were performed exactly according to Sambrook & Russell (2001) or Coligan et al. (1995).

4.3.1 SDS Polyacrylamide Gel Electrophoresis (SDS PAGE)

The glycine SDS PAGE was performed exactly like described in Sambrook & Russell (2001). For small proteins or peptides, however, the tricine SDS PAGE is better suitable, as it has better resolution in low molecular weight range. The tricine SDS PAGE was adapted from Schagger & von Jagow (1987).

Tricine SDS PAGE with urea:

Stock solutions:

1. buffer A

3M tris	181.71g/500ml
0.3% SDS	1.5g
0.05% NaN_3	0.25g
pH adjusted to 8.45 with HCl	

2. buffer B

acrylamide	24g/50ml
bis-acrylamide	0.75g

3. buffer C

acrylamide	23.25g/50ml
bis-acrylamide	1.5g

4. buffer D

   ammonium persulfate 10%

5. 6M urea 36.04g/100ml

The gels were prepared in chambers for 9 gels. Therefore the solutions volumes below are given always per 9 gels.

1. **stacking gel** (upper, poured as last, total volume 30ml)

   buffer A 7.5ml
   buffer B 2.5ml
   water 20ml
   buffer D 300µl
   TEMED 30µl

2. **spacer gel** (middle, poured as second, length 1cm, total volume 33ml)

   buffer A 10ml
   buffer B 6ml
   water 14ml
   6M urea 3ml
   buffer D 200µl
   TEMED 20µl

3. **separating gel** (lower, poured as first, length 4-5cm, total volume 62ml)

   buffer A 20ml
   buffer C 20ml
   water 13ml
   6M urea 9ml
   buffer D 400µl
   TEMED 40µl

Buffer D was prepared always freshly before use. Buffer D and TEMED were added immediately before the gels were poured. Spacer gel was poured immediately after pouring the separating gel so that they could mix together creating a polyacrylamide concentration gradient. Different voltage was used for distinct gels:

stacking gel: 25-30V
spacer gel: 50V
separating gel: 75-80V

Different buffers were used for anode and cathode:

1. anode buffer (+)

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

2. cathode buffer (-)

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

   0.1M tricine 17.9g/l

   0.1% SDS

4.3.2 Staining of Proteins

Staining of proteins with Coomassie-Blue and with Ponceau-Red was performed like described in Sambrook & Russell (2001). The silver staining, however, was little modificated.

Silver Staining:

Stock solutions:

1. solution 1

   300ml ethanol

   150ml acetic acid

   water up to 1000ml

2. solution 2

   41g sodium acetate

   250ml ethanol

   water up to 1000ml

   immediately before use add: 0.1g/50ml sodium thiosulfate

   250µl/50ml glutardialdehyde

3. solution 3

   1g silver nitrate

   water up to 1000ml

   immediately before use add: 15µl/50ml formaldehyde

4. solution 4

   25g sodium carbonate

   water up to 1000ml

   pH adjusted to 11.5 with carbonate/hydrogencarbonate

   immediately before use add: 15µl/50ml formaldehyde

5. solution 5

   18.6g EDTA

   water up to 1000ml

The gels were stained in the following manner:

solution	washing time
1	1h
2	1-12h (overnight)
water	3x10min
3	30min
4	-
5	-

4.3.3 Determination of Protein Concentration

The concentration of proteins in solutions was estimated with the assistance of the Bradford reagent (BioRad; Bradford, 1976). 10µl of the protein solution (or 1µl, if the protein solution is very concentrated) to be measured were added to 625µl of BioRad-reagent working solution (working solution was prepared by 1:5 dilution of BioRad-reagent stock solution in PBS buffer or water, stored in the fridge). Then the mixture was diluted with 400µl water. After thoroughly mixing the sample, the OD_{595} was measured. As a reference similar mixture was prepared with 10µl water instead of protein solution. OD was subsequently converted into the protein concentration on the basis of a BSA calibration curve.

4.4 NMR Samples Preparation

If not otherwise indicated, the samples for NMR spectroscopy were concentrated and dialyzed against PBS buffer. Typically, the sample concentration varied from 0.3 to 1.0 mM. Before measuring, the sample was centrifuged in order to sediment aggregates and other macroscopic particles. 450µl of the protein solution were mixed with 50µl of D_2O (5-10%) and transferred to an NMR sample tube.

4.5 Crystallization Procedure

Protein samples for crystallization trials were prepared in the following manner. The more than 95% pure protein sample was concentrated and purified by gel filtration (HiLoad Superdex column S75, S30 or S200) in low salt crystallization buffer. Collected fractions were pooled and concentrated using Amicon concentrating cell until the expected protein concentration was reached (5-100mg/ml). The membrane was washed several times in crystallization buffer before use. Subsequently, the sample was filtered through the 0.22µm Millipore filter, equilibrated previously twice with crystallization buffer. The sample was kept on ice or at RT. Reservoirs were filled with 500µl of the reservoir buffer. The hanging or sitting drop techniques were employed (see chapter 3.1).

5 Results and Discussion

5.1 Preliminary Investigations of E2F Protein Family

5.1.1 Expression and Solubility Tests of E2F Constructs in E.coli

Expression Test –Time Course

50ml LB-medium, containing appropriate antibiotika, were inoculated with a single bacterial colony from a fresh LB agar plate, and incubated at 37°C in a 100ml flask with vigorous shaking (280rpm). OD_{600} was monitored until the value 0.6-0.7 was reached. At that time (t = 0) the culture was induced with 1mM IPTG (endconcentration). The culture was grown overnight. 1ml samples for electrophoresis were taken before induction (t = 0), 1, 2, 3, 4, 5 hours after induction (t = 1, 2, 3, 4, 5), and after overnight culture incubation (t = N). The samples were centrifuged and pellet was dissolved in 50µl of the 2x SDS PAGE loading buffer (Sambrook & Russell 2001) and heated for 5 minutes. 15µl from every sample was loaded onto the gel. An example of a gel is show in Figure 5.1.1. Results were collected in Table 5.1.1.

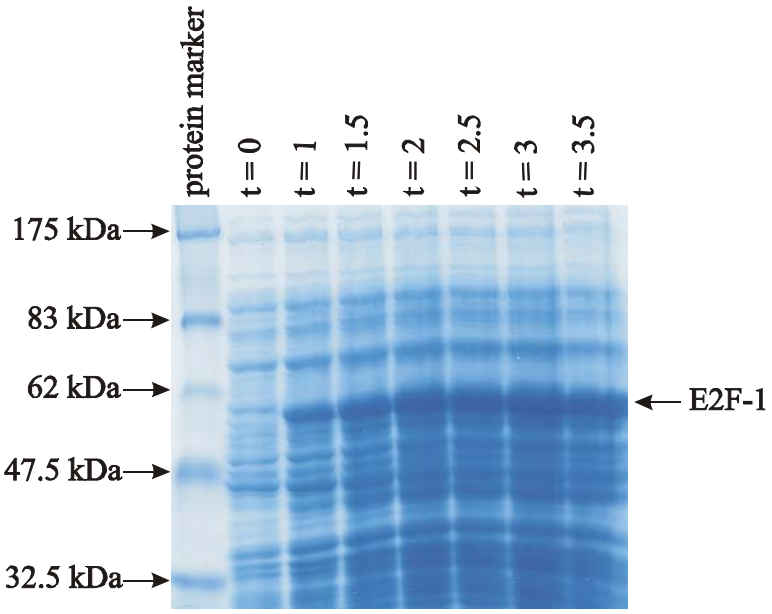


Figure 5.1.1. *Expression test of the His_6-tagged full length E2F-1. Coomassie stained SDS PAGE. The electrophoresis samples were taken every half an hour. For details refer to the text.*

Solubility Test

A bacterial culture with a tested construct was grown exactly like for expression test. After 4 hours of incubation after induction, culture was centrifuged for 20min at 4°C with 6000G. The pellet was resuspended in 5ml PBS buffer, and twice frozen in liquid nitrogen and thawed to disrupt the cells. Additionally, to ensure cells disruption, the suspension was shortly sonicated with a maximum sonicator power for 2x10s. The suspension was then centrifuged for 20min at 4°C with 12000G. The pellet was dissolved in 5ml buffer containing 8M urea, 0.1M NaH_2PO_4, 0.01M tris-HCl, pH 8.0 with 10mM β-ME. 20µl samples for electrophoresis were taken from supernatant as well as from the dissolved pellet. 20µl samples were mixed with 20µl of the 2x SDS PAGE loading buffer (Sambrook & Russell 2001) and heated for 5 minutes. 15µl from every sample was loaded onto the gel. An example of a gel is show in Figure 5.1.2. Results were collected in Table 5.1.1.

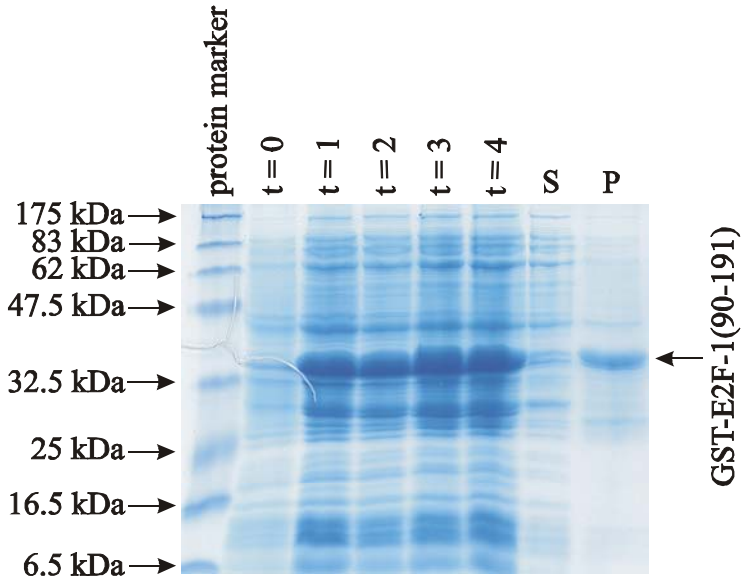


Figure 5.1.2. *Expression and solubility test of the GST-fused E2F-1 fragment (amino acids 90-191). Coomassie stained SDS PAGE. S stands for soluble part of the cell lysate. P for the insoluble part.*

All E2F constructs for expression in E.coli, which were tested, gave high or very high expression. They were, however, insoluble creating inclusion bodies. In this case, there were two possible attempts to the project, either finding other expression conditions, under which the expressed protein is soluble (see next section); or developing a refolding protocol to renature the insoluble protein (see chapter 5.1.3).

Table 5.1.1. *Results of the expression and solubility tests of the* E.coli *E2F constructs. For details refer to the text.*

#	construct	expression	solubility
1	GST-DP1	high	insoluble
2	GST-E2F1	high	insoluble
4	His_6-E2F1	very high	insoluble
5	tubulin-E2F1	low	insoluble
6	GST-E2F1 (90-437)	high	insoluble
9	GST-E2F1 (90-191)	high	insoluble
10	GST-E2F1 (90-238)	high	insoluble
1 & 4	GST-DP1 & His_6-E2F1	high	insoluble

Solubility Optimization Test

Eukaryotic proteins that are overexpressed in *E.coli* are very often insoluble creating so-called inclusion bodies. This is connected to the lost of protein tertiary structure and consequently to the lost of the protein activity. Because I was trying to find out the protein expression conditions under which the resulted protein is biologically active, I tried to find such conditions for protein expression that could result in the soluble protein, which is a hallmark of protein proper fold and biological activity. There are few factors that can directly influence the solubility of the protein produced in *E.coli*. They include temperature of the culture during the expression (24-37°C); optical density at which the culture was induced (OD_{600} = 0.5-1.0); the inductor (here: IPTG) concentration which was used for induction (0.05-2mM, endconcentration); as well as time after induction after which the culture was harvested (2-16 hours). More detailed introduction to the overexpression and purification of eukaryotic proteins in *E.coli* can be found e.g. in Marston (1986).
Cultures of bacteria containing tested construct for protein expression were grown similar to the cultures described in the solubility test. For every tested construct, however, two temperatures were tested (30 and 37°C), other varied parameters are presented in Table 5.1.2. Altogether for every given construct 18 different expression conditions were tested.

Two constructs were chosen for the optimization of the conditions for the protein expression to get possibly soluble protein. For the construct #6 (GST-fused E2F1 fragment, amino acids 90-437) all the conditions showed in Table 5.1.2 were tested. The cultures resulted in high protein expression, the protein was, however, always totally insoluble. Construct #9 (GST-fused E2F1 fragment, amino acids 90-191) resulted also mostly in

insoluble protein, although a conditions set was found under which the expressed protein showed to be at least partially soluble. The culture was supposed to be grown at 28°C, induced at OD_{600} = 0.7 with 0.1mM IPTG (endconcentration), and harvested after ca. 2.5 hours. Further studies with the construct are described in the following section.

Table 5.1.2. *Conditions tested for the optimization of the expressed protein solubility. All given sets of parameters were tested both for 30°C and 37°C.*

parameters set number	**culture induced at OD_{600}**	**induction with IPTG endconcentration [mM]**	**time from induction to harvest [h]**
1	**0.6**	0.5	2.5
2	**0.75**	0.5	2.5
3	**1.0**	0.5	2.5
4	0.75	**0.1**	2.5
5	0.75	**0.5**	2.5
6	0.75	**1.0**	2.5
7	0.75	0.5	**1**
8	0.75	0.5	**2.5**
9	0.75	0.5	**4**

5.1.2 GST-fused E2F-1 Fragment (Amino Acids 90-191)

The overexpressed polypeptide was purified with Glutathione Sepharose (Pharmacia, FRG). A bacterial pellet from 1 liter culture was resuspended in 30ml PBS, and 1ml 5% lysozyme, as well as traces of DNaseI, RNaseA, and $MgCl_2$ were added. After 1 hour incubation on ice, the sonication of the suspension followed, using microtip 4x2min output control 7, 50%. The mixture was then centrifuged for 10min at 4°C with 12000G. After filtration through the 0.45µm filter (Millipore, FRG), supernatant was loaded onto a 3ml Glutathione Sepharose column previously equilibrated in PBS. The protein was eluted with a gradient to GEB buffer.

Glutathione Elution Buffer (GEB):

10mM reduced glutathione

50mM tris-HCl, pH 8.0

0.05% NaN_3

The electrophoresis of the collected fractions (Figures 5.1.3 and 5.1.4) reveals that the eluted mixture of the proteins contains not only GST-fused E2F-1 fragment, but also GST alone and some other bands, probably GST with E2F-1 fragment degraded by proteases. This can mean that either E2F-1 fragment is very sensitive to proteases or it is rather unfolded. Further trials to work steadily at 4°C and with addition of proteases inhibitors set (Complete, Roche, FRG) did not bring any changes.

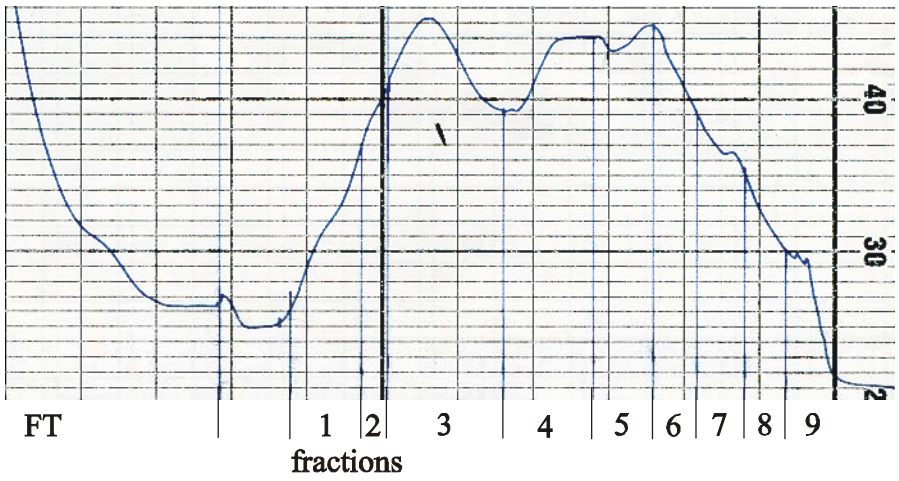


Figure 5.1.3. *The elution profile of the GST-fused E2F-1 fragment out of the Glutathione Sepharose. For details refer to the text.*

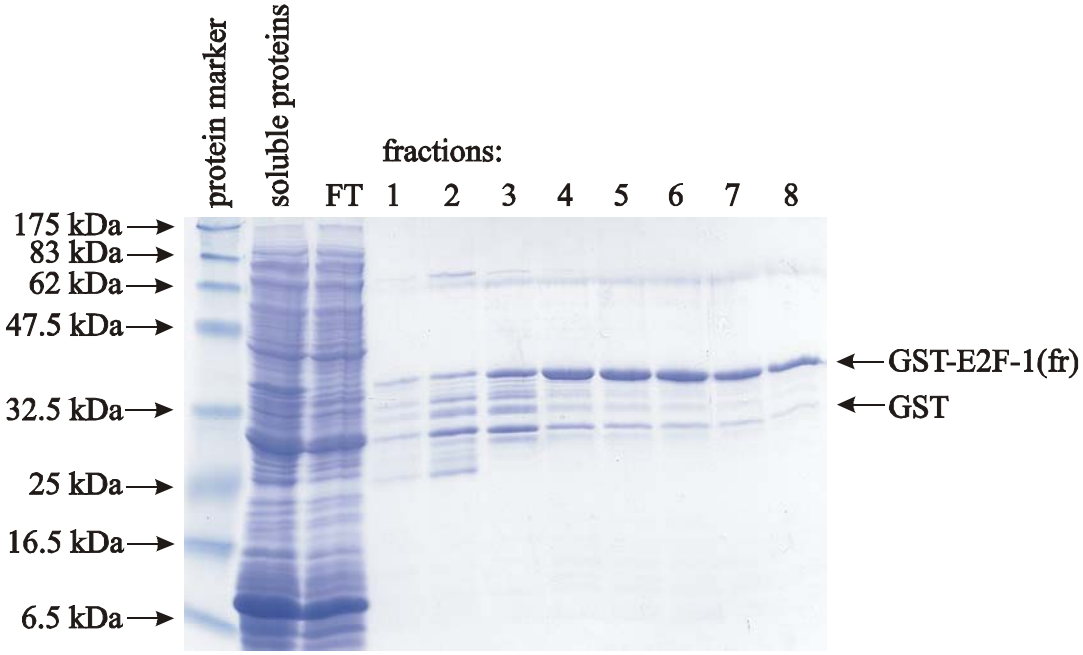


Figure 5.1.4. *SDS PAGE of the fractions eluted from Glutathione Sepharose. Compare to Figure 5.1.3.*

The fusion protein consisted in almost 85% of GST. To continue the work on the small E2F-1 fragment a specific proteolytical cleavage was needed. The construct contained a specific cleavage site for factor Xa protease. The digestion was performed exactly according to the suggestions of the producer (NEB, FRG). All fractions eluted from Glutathion Sepharose, which contained the fusion protein were pooled and four similar digestion mixtures were prepared. Two of them were supposed to be the controls and did not contain contained factor Xa. All the mixtures were incubated at 37°C for 2 or 16 hours. Samples for electrophoresis were taken (Figure 5.1.5).

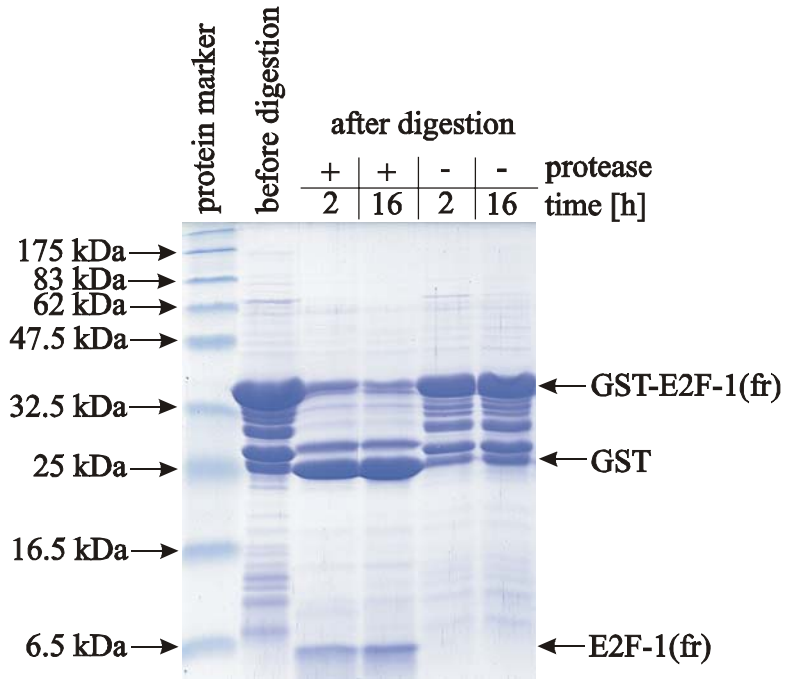


Figure 5.1.5. *Results of the proteolytical cleavage with factor Xa. Coomassie stained SDS PAGE. For details refer to the text.*

The mixtures without addition of the protease did not change with the tested incubation time, which means that the fusion protein degradation is not continued or is very slow. This can also mean that the additional bands on the gel are not due to the proteases but because of the additional start or stop points for the transcription. These were, however, not found by the sequence analysis. The digestion mixtures contain apparently the free GST protein as well as the digested small E2F-1 fragment. The fragment, however, was totally insoluble. A few additional trials were performed including changes in protein preparation like using batch purification, proteolytical cleavage on the Glutathione Sepharose column or on SP Sepharose column but they did not bring any changes. The precipitated E2F-1 fragment was dissolved in

a buffer containing 8M urea and 10mM β-ME (compare the solubility test) and subsequently dialysed against PBS but the protein precipitated again.

5.1.3 His-tagged Full Length E2F-1 Construct

The full-length histidine-tagged E2F-1 was shown (previous section, Figure 5.1.1) to give a very high overexpression. The resulted polypeptide was, however, fully insoluble. Therefore a few trials were performed to establish a refolding protocol for the protein. The inclusion bodies were purified (see below) and solubilized in a buffer containing either 8M urea or 6M guanidine hydrochloride supplemented with 10mM β-ME. A denaturant was then removed using many different protocols to allow protein folding. The most important parameters sets and results are collected in Table 5.1.2.

There was no logical explanation found, why some refolding procedures were more successful than others. Full-length E2F-1 protein contains 6 cysteine residues, which let us suppose that the protein aggregation is mainly due to the creation of the intra- and intermolecular disulfide bridges. This assumption was, however, falsified by the “dilution to PBS”-procedure. The method gave the same result independent on the presence of the reducing agent. Another reasonable explanation would be that the guanidine hydrochloride as a denaturant is in this case less efficient. To falsify, however, which parameter is really responsible for the protein aggregation or solubility, more detailed studies have to be performed.

short name or reference	**inclusion bodies solubilization buffer**	**refolding procedure**	**refolding buffer**	**resulted protein**	**spectrum file**
Nevin's refolding (Yee et al., 1991)	6M GdnHCl 50mM tris, pH 7.9 12.5mM $MgCl_2$ 1mM EDTA 1mM DTT 20% glycerol 100mM KCl 0.1% NP-40, overnight at RT	sample in solubiliz. buf. loaded onto the PD10 desalting column (Pharmacia) pre-equilibrated in refold. buf., at RT	50mM tris, pH 7.9 12.5mM $MgCl_2$ 1mM EDTA 10mM β-ME 20% glycerol 100mM KCl 0.1% NP-40	insoluble	---
MDM2 protocol (see section 5.2.1)	6M GdnHCl 100mM tris, pH 8.0 1mM EDTA 10mM DTT	after solubilization pH changed to 3.0, dialysis against 4M GdnHCl, pH 3.5 incl. 10mM DTT; dilution in several pulses to the refold. buf., left 16 h at RT	10mM tris, pH 7.0 1mM EDTA 10mM DTT	insoluble	---
alkaline refolding (Marston et al., 1984)	8M urea 50mM tris, pH 8.0 1mM EDTA 50mM NaCl no red., 1h at RT	10-fold dilution to the refold. buf., 30 min incubation at RT; pH adjusted back to 8.0, concentrated, dialysed	50mM KH_2PO_4, pH 10.7 150mM NaCl 0.05% NaN_3 no red.	soluble	rbE2F_990108
dilution to PBS - red.; (similar to alkaline)	8M urea 50mM tris, pH 8.0 1mM EDTA 50mM NaCl no red., 1h at RT	10-fold dilution to the refold. buf., 30 min incubation at RT; concentrated, dialysed	50mM KH_2PO_4, pH 8.0 150mM NaCl 0.05% NaN_3 no red.	soluble	mpH6_990114
dilution to PBS + red.	8M urea 50mM tris, pH 8.0 1mM EDTA 50mM NaCl 10mM DTT overnight at RT	10-fold dilution to the refold. buf., 30 min incubation at RT; concentrated, dialysed	50mM KH_2PO_4, pH 8.0 150mM NaCl 0.05% NaN_3 10mM DTT	soluble	see all files in Table 5.1.3

Table 5.1.2. *Refolding procedures tested with the full-length E2F-1.(red. = reducing agent)*

Parallel to the refolding tests, another trials were performed aimed to purify the E2F-1 protein. Following protocol for the purification and refolding of the histidine-tagged full-length E2F-1 was established. The cell pellet from the 1 liter bacterial culture (described in section 4.2.2) was resuspended in PBS buffer, treated with lysozyme and sonicated like described in the previous section. Low speed centrifugation with 12000G for only 6 minutes at 4°C followed. Resulted inclusion bodies pellet was washed twice with PBS buffer containing 0.5% (v/v) Triton X-100 every time with the subsequent low speed centrifugation, solubilized in a buffer containing 8M urea and 10mM β-ME (compare to the section 5.1.1, solubility test), and centrifuged for 30 minutes at RT with at least 18000G. The supernatant was loaded onto a 10ml Q Sepharose (Pharmacia, FRG) column equilibrated previously in the solubilization buffer. The protein was eluted with a gradient to the solubilization buffer supplied with 1M NaCl. The denaturant was removed by the mentioned above "alkaline-refolding" or by rapid dilution to the 10-fold volume of the PBS buffer containing (or not) 10mM DTT, and subsequently dialysed against PBS buffer with addition of 5mM DTT. The three mentioned refolding procedures resulted in the soluble E2F-1 protein.

The protocol described above was used to produce E2F-1 samples for NMR spectroscopy including ^{15}N-uniformly labeled E2F-1 sample. Performed NMR measurements, their data files and results are collected in the Table 5.1.3.

Table 5.1.3. *NMR experiments with full-length E2F-1. Binding studies were performed in equimolar concentrations. RB protein (C-terminal 56 kDa fragment) was a kindly gift from firma Roche (Penzberg, FRG).*

sample	**exp. type**	**data file**	**results**
^{15}N E2F-1	1D, HSQC	msE2F1N15_0324 rpE2F1N15_0324 rpE2F15N_00404-1,2	protein is unfolded
^{15}N MDM2 + non labeled E2F-1	HSQC	rpMDME2F1_0403	no change
^{15}N E2F-1 + non labeled MDM2	HSQC	rpE2F15N_00404-3,4	no change
^{15}N E2F-1 + RB	1D	msE2FRB_0425	sample precipitated
^{15}N E2F-1 titration with TFE	HSQC	msE2FTFE_0425	no change

NMR spectra revealed that E2F-1, in the prepared samples, has not a defined three-dimensional structure (Figures 5.1.6 and 5.1.7). One-dimensional (1D) spectrum of E2F-1 (Figure 5.1.6) shows clearly that protein is not folded. Both the N-H region of the spectrum

(around 8 ppm, Figure 5.1.6.B) and the C-H region (around 1 ppm, Figure 5.1.6.C) contain very broad peaks and not sharp and well dispersed like in the case of the structured protein. The HSQC spectrum of the ^{15}N-uniformly labeled E2F-1 (Figure 5.1.7) shows that all peaks in the spectrum (with exception of peaks coming from the side chains that have a stable position at around 6.7 and 7.5 ppm) are concentrated around 8.3 ppm, which is an average value for the random coil protein conformation (compare to a spectrum of a properly folded protein: Figure 5.2.2).

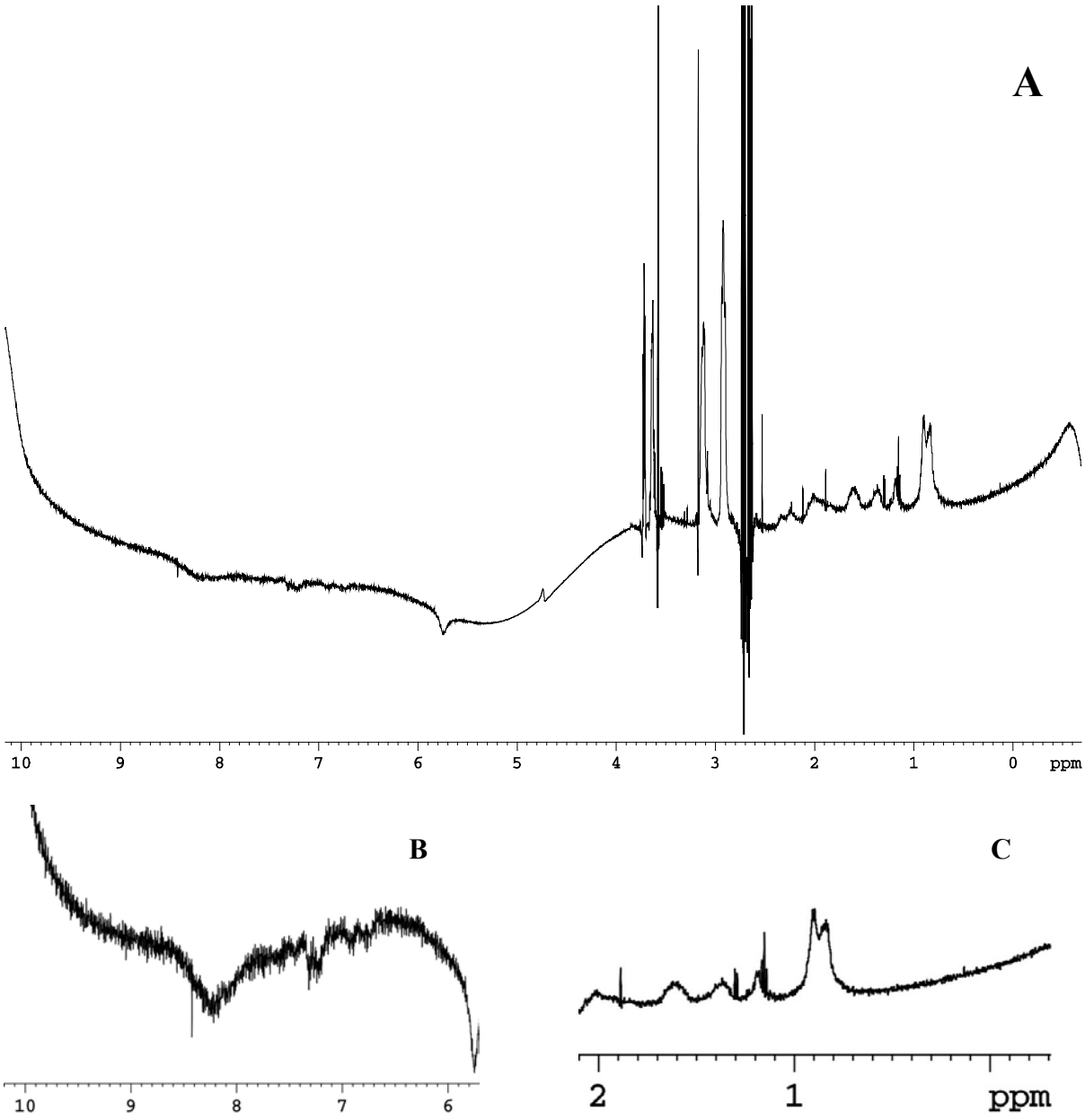


Figure 5.1.6. A. *1D spectrum of the E2F-1. Two fragments of the spectrum that give information about protein folding were separately phased and zoomed to make results more visible.* B. *N-H region.* C. *C-H region. For details refer to the text.*

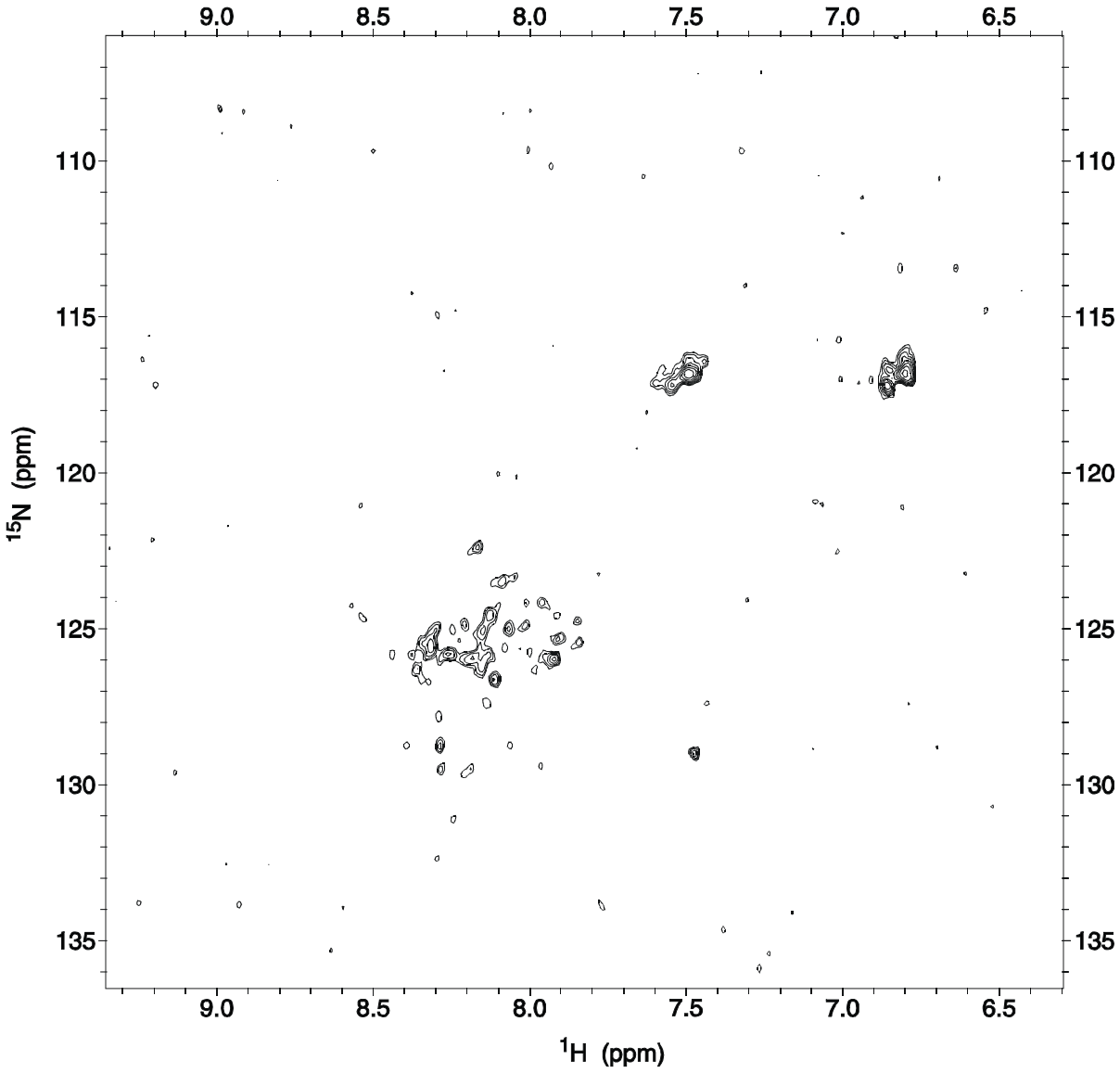


Figure 5.1.7. *HSQC spectrum of the ^{15}N-labeled E2F-1.*

Additionally, a few experiments more were done (Table 5.1.3), assuming that E2F-1 binding partners (MDM2, RB) can induce its folding. The spectrum of ^{15}N E2F-1, however, did not show any difference after addition of MDM2. Similarly, E2F-1 addition to the ^{15}N MDM2 sample did not indicate any specific binding in HSQC spectrum of MDM2. RB addition to the ^{15}N E2F-1 protein solution caused precipitation of both proteins, which is not explained. Trifluoroethanol (TFE) that is known to induce the secondary structure in proteins (Luo et al., 1997; Arunkumar et al., 1997) was also unsuccessfully used to induce any structural changes in E2F-1.

5.1.4 BEVS Constructs Expression Tests

All E2F constructs for expression in *E.coli*, which were tested, resulted in an insoluble protein. Therefore it was reasonable to test another expression system to check if it will bring any change in protein solubility and possibly produce a folded and biologically active protein. The **b**aculovirus **e**xpression **v**ectors **s**ystem (BEVS) has many advantages in comparison to bacterial system. Insect cells, which are employed here as the host cells, create a eukaryotic environment for protein expression raising the possibility of the proper protein folding. Additionally this system is capable of performing several post-translational modifications (N- and O-linked glycosylation, phosphorylation, acylation, amidation, carboxymethylation, isoprenylation, signal peptide cleavage and proteolytic cleavage). The sites where these modifications occur are often identical to those of the authentic protein in its native cellular environment. For details about the BEVS see the manual by O'Reilly et al. (1994). All protocols for the work with the system were taken from the manual.

Four baculoviral constructs for protein overexpression in insect cells were tested (Table 5.1.4). The expression tests were done as described in O'Reilly et al. (1994). The SF (*Spodoptera frugiperda*) cells were infected at the concentration of 2mln. cells/ml with a high-titer baculovirus stock. 1ml samples for SDS PAGE were taken after 0, 24, 37, 49, 54 hours **p**ost **i**nfection (pi) and prepared like the samples in the expression tests in *E.coli* (see section 5.1.1) with the only difference that they were heated for at least 20 minutes. Results were collected in Table 5.1.4. An example of the expression test is shown in Figure 5.1.8 (lanes marked as "DP-1").

Table 5.1.4. *Results of the E2F expression tests in BEVS.*

construct	**expression**
hemagglutinin(HA)-tagged full length E2F-1	very low
hemagglutinin(HA)-tagged full length E2F-4	low
hemagglutinin(HA)-tagged full length DP-1	low
not tagged full length DP-2	very low

The protein overexpression is low in comparison to *E.coli* constructs, however still relatively high, the percentage of the overexpressed protein between the huge number of baculoviral and insect-cellular proteins is, however, so low that western blots have to be done to notice the overexpression. The overexpression level is very promising, it would be, however, difficult to purify the proteins of interest out of the cell lysate which contains enormous number of other proteins.

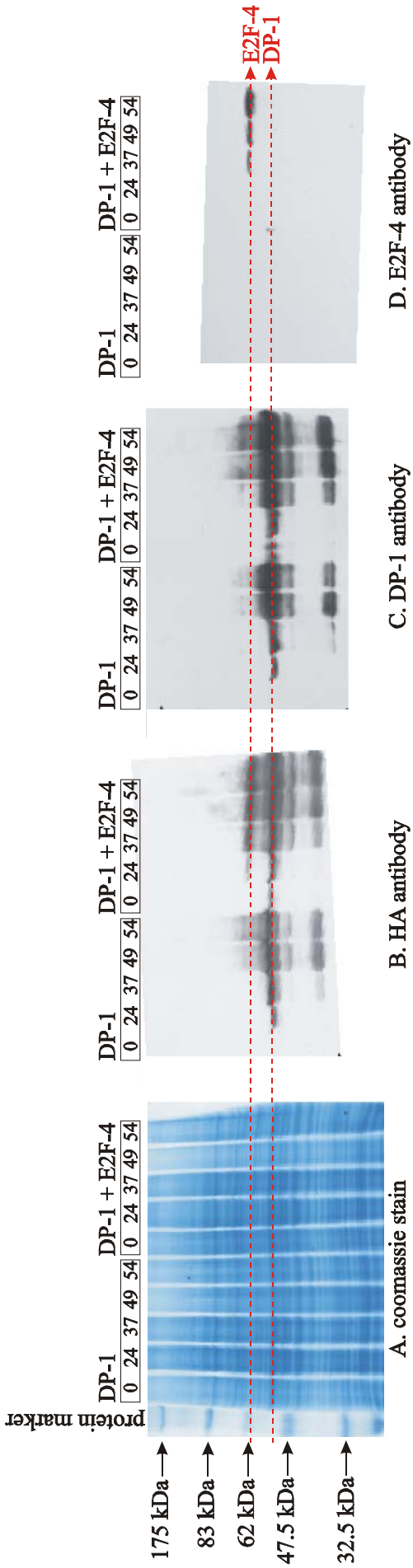


Figure 5.1.8. *Expression test of the full length hemagglutinin(HA)-tagged DP-1 in baculovirus expression vectors system (lanes marked as "DP-1" on every gel) as well as HA-DP-1 and HA-E2F-4 co-expression experiment (lanes marked as "DP-1 + E2F-4") are shown.* A. *Coomassie stained SDS PAGE gel.* B. C. D. *Western blots of the same gel using antibodies against hemagglutinin, DP-1, and E2F-4 respectively. The numbers above the lanes give time post infection (in hours) after which the samples for electrophoresis were prepared.*

Therefore further attempts to purify the protein expressed in insect cells should include cloning of the protein c-DNA into the baculoviral vector together with a tag (e.g. His_6-tag) or as a fusion protein (e.g. GST-fusion). Another possibility would be to prepare an affinity column containing antibodies against hemagglutinin-tag (cheaper version) or against particular E2F-family members, bound to agarose. Both possibilities could make the baculoviral construct useful for protein purification for structural studies, which require a huge amount of the material as well as high degree of purity (more than 95%).

Using the same constructs the co-expression trial was also performed. Insect cells culture was infected simultaneously with both DP-1 and E2F-4 overexpressing baculoviruses. The same protocol like for the expression tests was used. Results are shown in Figure 5.1.8 (lanes marked as "DP-1 + E2F-4"). The overexpression of both proteins is evidenced by using three distinct antibodies. The straight-forward co-expression of the proteins that are able to dimerize

is an additional bonus of the BEVS, as it raises the possibility of producing properly structured proteins. To make an advantage from the protein co-expression in the BEVS, however, the same considerations as made for the expression test results, should be taken into account.

5.2 Chalcone Derivatives Are Inhibitors of MDM2 and p53 Interactions

5.2.1 Protein Expression, Refolding and Purification

The recombinant human MDM2 protein was obtained from an *Escherichia coli* BL21 expression system and contained the first 118 N-terminal residues of human MDM2 cloned in a pQE-40 vector (Qiagen), C-terminally extended by an additional serine residue. Inclusion bodies were washed twice with the PBS buffer containing 0.05% Triton X-100 with subsequent low-speed centrifugation (12000G), and solubilized with 6M guanidine hydrochloride in 100mM tris-HCl, pH 8.0, including 1mM EDTA and 10mM DTT (10ml buffer per 1g inclusion bodies). After lowering pH to 3-4, the protein was dialyzed at 4°C, against 4M guanidine hydrochloride, pH 3.5, including 10mM DTT, until equilibrium was reached. For renaturation the protein was diluted (1:100) into 10mM tris-HCl, pH 7.0 including 1mM EDTA and 10mM DTT by adding the protein in several pulses. Refolding was performed for overnight at room temperature. Ammonium sulfate was added to a final concentration of 1M and the refolded human MDM2 was applied to hydrophobic interaction chromatography (batch purification) using Buthyl Sepharose 4 Fast Flow (Pharmacia, FRG). Because of the low binding capacity of the medium, 100ml bead volume per 1liter bacterial culture was used. The protein was eluted with 0.1M tris-HCl, pH 7.2 supplied with 5mM DTT. Finally, all fractions containing MDM2 were pooled, concentrated, and applied to a HiLoad 26/60 Superdex 75pg gel filtration column (Pharmacia, FRG). The running buffer contained 50mM KH_2PO_4, 50mM Na_2HPO_4, pH 7.4, 150mM NaCl, 5mM DTT, 0.02% NaN_3, and protease inhibitors (Complete^{TM}, Roche, FRG). All fractions with monomeric human MDM2 were pooled and concentrated with Amicon concentrating cell (cut off 10kDa) up to 1mM for NMR spectroscopy. The protein was stored at –20°C.

The p53 peptide comprising residues E17 to N29 of human p53 was chemically synthesized and contained an additional cysteine at the N-terminus. The peptide was purified by reversed phase chromatography.

5.2.2 p53/MDM2 Binding ELISA

Interference of the p53/MDM2 binding by low molecular weight compounds was measured in a 96-well polypropylene round-bottom microtiter plate (Costar, Serocluster). Human MDM2 (amino acids 1-118, at 40 nM) was preincubated with PBS/0.05% Tween50 (PBST)/10% DMSO or low molecular weight compounds. After 15-min incubation of the

sample, 100 nM of a p53-derived peptide (MPRFMDYWEDL, biotinylated, synthesized on solid phase) was added (Böttger et al., 1996). As a negative control, buffer only was added into separate wells (blanks). After another 30 min, the incubation mixture was added to 96-well plates coated with streptavidin. After 1 h, the wells were extensively washed with PBS/0.05% Tween50. Then, the MDM2 specific antibody (N20, Santa Cruz Biotechnology) in PBST and 1% casein was added. After 1 h, the wells were thoroughly washed with PBST and the secondary antibody (anti-rabbit-IgG-POD, RMB) in PBST and 1% casein was added. After another hour, the wells were washed with PBST, and the peroxidase substrate ABTS was added. After 30 min, OD405/490 nm was determined with a Dynatech MR 7000 ELISA reader. Calculation of inhibition was done as follows: [1 - (OD405/490 nm compounds - OD405/490 nm blanks)/(OD405/490 nm 100% values - OD405/490 nm blanks)] × 100 = % inhibition. Compounds were titrated to determine IC_{50} values twice (range of compound concentration: 0.5, 1, 5, 10, 25, 50, 125, 250 µM).

5.2.3 Gel Shift Assay

The DNA binding assay was performed using active fractions of human p53 protein expressed in baculovirus-infected insect cells and purified on Hi-Trap Heparin-Sepharose (Pharmacia Biotech) in a linear gradient from 0.1 to 0.85 M KCl (Hansen et al., 1996). MDM2 containing the first 118 amino acids was cloned as a GFP (green fluorescent protein) fusion protein at its N-terminus, which served to enlarge the protein to obtain a significant shift in the electrophoretic gel mobility shift assay (EMSA). (Larger fragments of the MDM2 protein tended to aggregate and were therefore not used.) Additionally, MDM2 was His-tagged C-terminally by $(His)_6$, which were added via a linker segment containing (Ser-Arg-Gly-Ser) for convenient purification. The construct was cloned in a modified pQE-40 vector (Qiagen) and expressed in E. coli BL21 (DE3) at 22°C as soluble protein. The lysate was purified using a Talon column (Clontech) according to standard protocols. p53 was bound to its specific, double-stranded DNA consensus site (PG) (El-Deiry et al., 1992), which was labeled with [γ-^{32}P]ATP. To ensure sequence-specific binding, a 20-fold (200 ng) excess of nonlabeled supercoiled competitor DNA (pBluescript II SK+, Stratagene) was included. p53 protein was used at a concentration of 200 nM and MDM2 at 2 µM. Despite the apparent excess of the proteins over the used DNA, the active fraction of the total protein preparation is so small that the DNA is still in excess, as can also be seen from the free DNA in Figure 5.2.1.

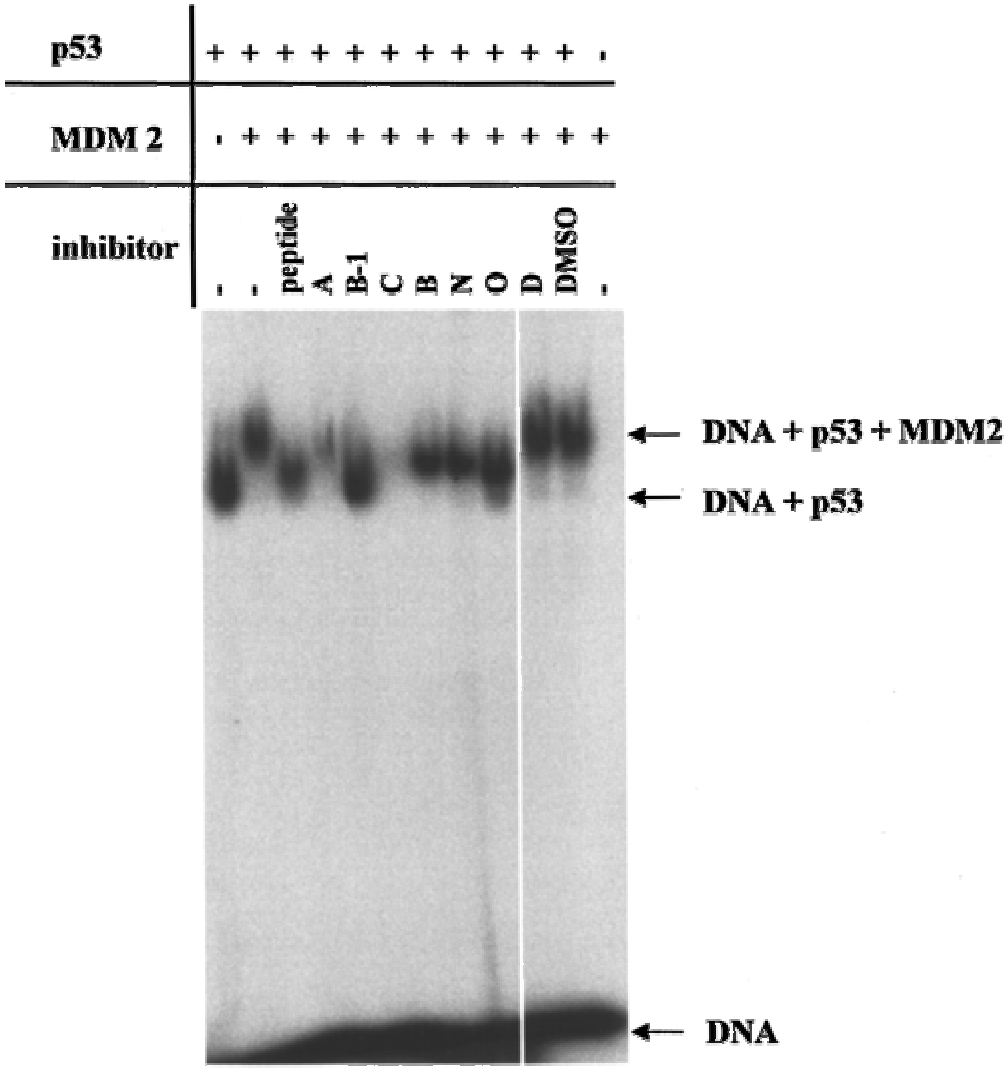


Figure 5.2.1. *Effect of chalcones on DNA binding activity of p53/MDM2 complexes. Human p53 was analyzed for DNA binding following incubation with and without MDM2 using a electrophoretic gel mobility shift assay (EMSA). Preformed p53/MDM2 complexes were subjected to incubation with competing p53 peptide (250 µM) or low molecular weight compounds (1 mM). Since the compounds contained a final concentration of 5% DMSO, a DMSO control of 5% was included.*

The p53 peptide was used at 250 µM, compounds at 1 mM. p53 was preincubated with MDM2 at RT for 30 min prior to addition of compounds for 30 min at 4°C and, finally, addition of DNA in DNA binding buffer for another 15 min at 4°C. The DNA binding buffer contained: 20% (v/v) glycerol, 50 mM KCl, 40 mM Hepes, pH 8, 5 mM DTT, 0.1% Triton X-100, 10 mM $MgCl_2$, 1.0 mg mL^{-1} bovine serum albumin. The reaction mix was loaded onto a 4% native polyacrylamide gel and separated at 200 V for 2 h at 4 C. The gel was dried and the DNA was detected by autoradiography.

5.2.4 NMR Spectra and Assignments

All NMR spectra were acquired at 290 and 300 K on Bruker AMX500, DRX500, DRX600, and DMX750 spectrometers. Typically, NMR samples contained up to 0.5 mM of protein in 50 mM KH_2PO_4, 50 mM Na_2HPO_4, 150 mM NaCl, pH 7.4, 5 mM DTT, 0.02% NaN_3, and protease inhibitors. The quality of the spectra for MDM2 with and without inhibitors was reduced by aggregation, especially at concentrations higher than 0.5 mM at pH 7.4 and 300 K. Since concentrated samples remained stable for approximately 1 day, only highly sensitive experiments could be performed. A nearly complete assignment of the backbone $^1H^N$ and ^{15}N NMR resonances was obtained for the uncomplexed MDM2 (apo-MDM2; Figure 5.2.2).

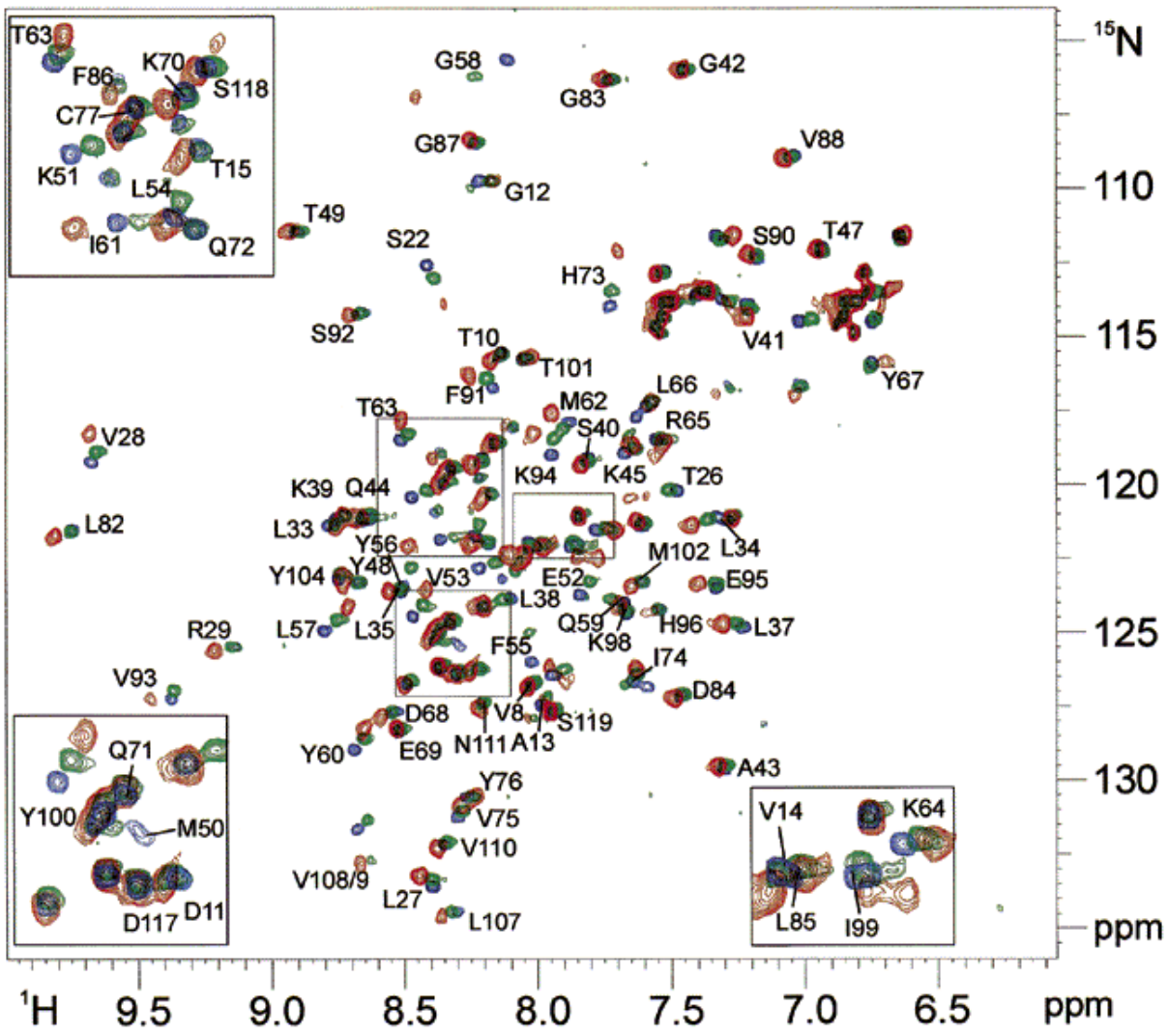


Figure 5.2.2. *500 MHz 2D ^1H-^{15}N HSQC spectrum of human MDM2 titrated with increasing amounts of chalcone C. Cross-peaks for apo-MDM2 are marked in blue; green and red cross-*

peaks indicate 50 and 100% complexation of MDM2 by chalcone C. Residue specific assignment of the backbone ^{1}H and ^{15}N frequencies is indicated.

Backbone sequential resonances were assigned with CT-HNCA, CBCA(CO)NH using the WATERGATE sequence, and in part with 2D TOCSY (mixing time of 42 ms), 2D NOESY (mixing time 120 ms), 3D ^{15}N-TOCSY-HSQC (spin-lock period of 36 ms), and 3D 15N-NOESY-HSQC (mixing time of 120 ms) experiments (Grzesiek & Bax, 1992, Jahnke et al., 1995), and by selective enrichment using ^{15}N-Leu, Phe, Val, and reverse ^{14}N-His samples of MDM2. ^{15}N-{^{1}H} heteronuclear NOE was measured using a modified version of the experiments as described previously (Farrow et al., 1994). NOE values were calculated by scaling ratios of peak heights in the NOE experiment with ^{1}H presaturation and the standard HSQC experiment obtained from the same sample. Recording of the NOE experiment without proton saturation using the same sample was not possible due to the fast precipitation of apo-MDM2 samples. This simplified approach introduces an additional error of approximately 10-20% to the NOE values. The experiment was recorded in an interleaved manner so that precipitation of the protein results in broadening of the signals but does not affect the extracted NOE values (Farrow et al., 1994).

5.2.5 Ligand Binding

All chalcone derivatives used in this study have been synthesized according to standard Claisen-Schmidt aldol condensation protocols as previously published (Daskiewicz et al., 1999, Bois et al., 1999). A total of 50 chalcone derivatives were synthesized (Figure 5.2.4). NMR measurements consisted of monitoring changes in chemical shifts and line widths of the backbone amide resonances of uniformly ^{15}N-enriched MDM2 samples (Shuker et al., 1996, McAlister et al., 1996) in a series of HSQC spectra as a function of a ligand concentration (Shuker et al., 1996). No changes in chemical shifts were observed between samples of different concentrations (0.03-0.5 mM) and pH values between 6.5 and 7.5. For titration experiments, 0.1-0.3 mM of human MDM2 in 50 mM KH_2PO_4, 50 mM Na_2HPO_4, 150 mM NaCl, pH 7.4, and 5 mM DTT was used. The chalcone derivatives were lyophilized and finally dissolved in DMSO-d_6. No shifts were observed in the presence of 1% DMSO (the maximum concentration of DMSO in all NMR experiments after addition of inhibitors). All chalcone-MDM2 complexes showed a continuous movement of several NMR peaks upon addition of increasing amounts of inhibitors. From these experiments, the spectra of MDM2

could be assigned unambiguously. The complexes of human MDM2 and the chalcones were prepared by mixing the protein and the ligand in the NMR tube. Typically, NMR spectra were recorded 15 min after mixing at room temperature. An initial screening of all compounds used in this study was performed with a 10-fold molar excess of chalcone to human MDM2. All subsequent titrations were carried out until no further shifts were observed in the spectra. Saturating conditions were achieved at a molar ratio of chalcone to MDM2 of 6 for chalcone A, of 2 for chalcone B, of 2 for chalcone B-1, and of 6 for chalcone C, for example. Typically, the concentration of human MDM2 was 0.1 mM and the final concentration of the chalcone ligand was 50 mM in each titration. All K_D values obtained by NMR spectroscopy are based on at least six data points. From the independently determined IC_{50} values and the K_D constants, one ligand binding site for these chalcones per MDM2 is calculated taking into account the molar ratio of ligand to protein in the NMR experiments. Quantitative analysis of induced chemical shifts were performed on the basis of spectra obtained at saturating conditions of each chalcone. Analysis of ligand-induced shifts was performed by applying the equation of Pythagoras to weighted chemical shifts: $\Delta\delta_c(^1H, ^{15}N) = [\{\Delta\delta(^1H)^2 + 0.2 \times \Delta\delta(^{15}N)^2\}^{0.5}]$. The p53 peptide/MDM2 complex was long-lived on the NMR chemical shift time scale (lifetimes >> 2 ms) (Wüthrich, 1986). Two separate sets of resonances were observed in the ^1H-^{15}N HSQC spectra, one corresponding to free MDM2 and the other to MDM2 bound to the p53 peptide. For well-resolved, isolated peaks, the assignment of Figure 3 could be transferred to the resonances in the peptide complex (54% of all backbone amide resonances in the ^1H-^{15}N HSQC). For the rest of the shifts, assignment of $\Delta\delta_c(^1H, ^{15}N)$ upon complex formation was carried out in a conservative manner, i.e., for these shifts the distance in ppm to the closest peak in complexed MDM2 was chosen. In addition, all selectively enriched samples of human MDM2 (^{15}N-Val, ^{15}N-Leu, ^{15}N-Phe, and reverse ^{14}N-His) were titrated with the p53 peptide to confirm a subset of MDM2/p53 complex assignments. Only $\Delta\delta_c(^1H, ^{15}N)$ values larger than 0.1 ppm were considered to be significant. $\Delta\delta_c(^1H, ^{15}N)$ smaller than 0.1 ppm were found for 37 residues. Erroneous conclusions could result if some of the residues with $\Delta\delta_c(^1H, ^{15}N) < 0.1$ ppm were actually in contact with the inhibitor. However, the internal consistency of our results corroborates our analysis; for example, no core buried residue was found that had $\Delta\delta_c(^1H, ^{15}N) > 0.1$ ppm. Furthermore, all residues of human MDM2 involved in binding to the p53 peptide also show significant shifts $\Delta\delta_c(^1H, ^{15}N)$ upon complexation with the peptide (Kussie et al., 1996). For compounds B and B-1 (Figure 5.2.3, panels C and D), the maximum shifts shown at $\Delta\delta_c = 0.5$ ppm correspond to the cross-peaks of the folded core of MDM2 whose line-widths broaden 2-fold upon addition of either B or B-

1 in the molar ratio of B-1 to MDM2 1:1 and disappear thereafter at the titration ratio 2:1 (McAlister et al., 1996).

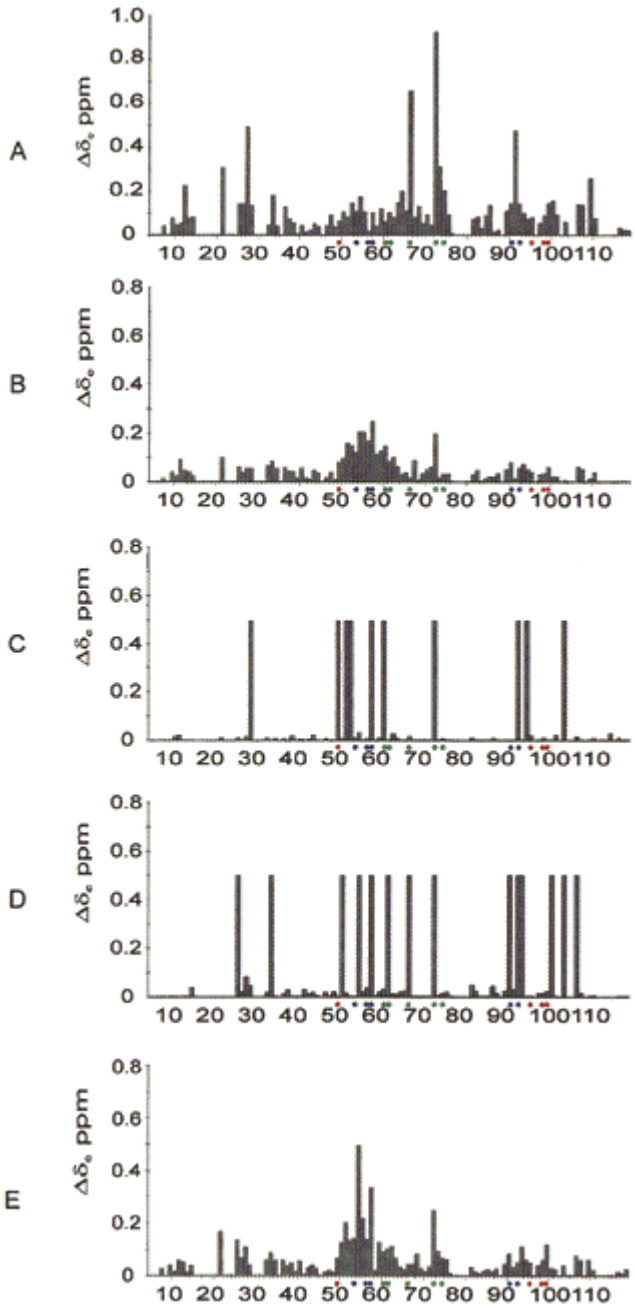


Figure 5.2.3. *Plots of induced differences in the NMR chemical shifts versus the amino acid sequence.* (A) *The p53 peptide;* (B) *inhibitor A;* (C) *inhibitor B;* (D) *inhibitor B-1 (for the maximum induced shifts for B and B-1 see explanation in experimental procedures);* (E) *inhibitor C. Red, blue, and green dots mark the leucin-, tryptophan-, and the phenylalanine-binding site on human MDM2 (refer to Figure 5.2.6).*

Compound D (Figure 5.2.4) was studied as a negative control because it did not inhibit MDM2 binding to a p53 peptide as measured by ELISA. This compound does not bind to apo-MDM2, as no ^{1}H and ^{15}N shifts greater than 0.1 ppm were observed in the NMR spectra. As this compound was available in our laboratory and because of its similar size as compared to the chalcone skeleton, we have selected this heterocyclic system as a negative control for any organic compound. Other negative control NMR titration experiments included the chemically synthesized chromophore of the green fluorescent protein as well as a synthetic 22-residue peptide. None of the control ligands led to significant chemical shift perturbations (data not shown). Chalcone B-1 generally enhances the intrinsic tendency of MDM2 to aggregate at higher concentrations. Therefore, an additional experiment was performed to test their specificity and to rule out a property as a general protein precipitant. For this purpose, the human tumor suppressor $p19^{INK4d}$ was purified as previously described (Baumgartner et al., 1998).

Chalcone B-1 did not induce aggregation of $p19^{INK4d}$ when applied under the same experimental conditions.

5.2.6 Chalcones Are MDM2 Antagonists

Derivatives of the chalcone class have been shown to inhibit MDM2 binding to a p53 peptide in a two-site ELISA (Figure 5.2.4).

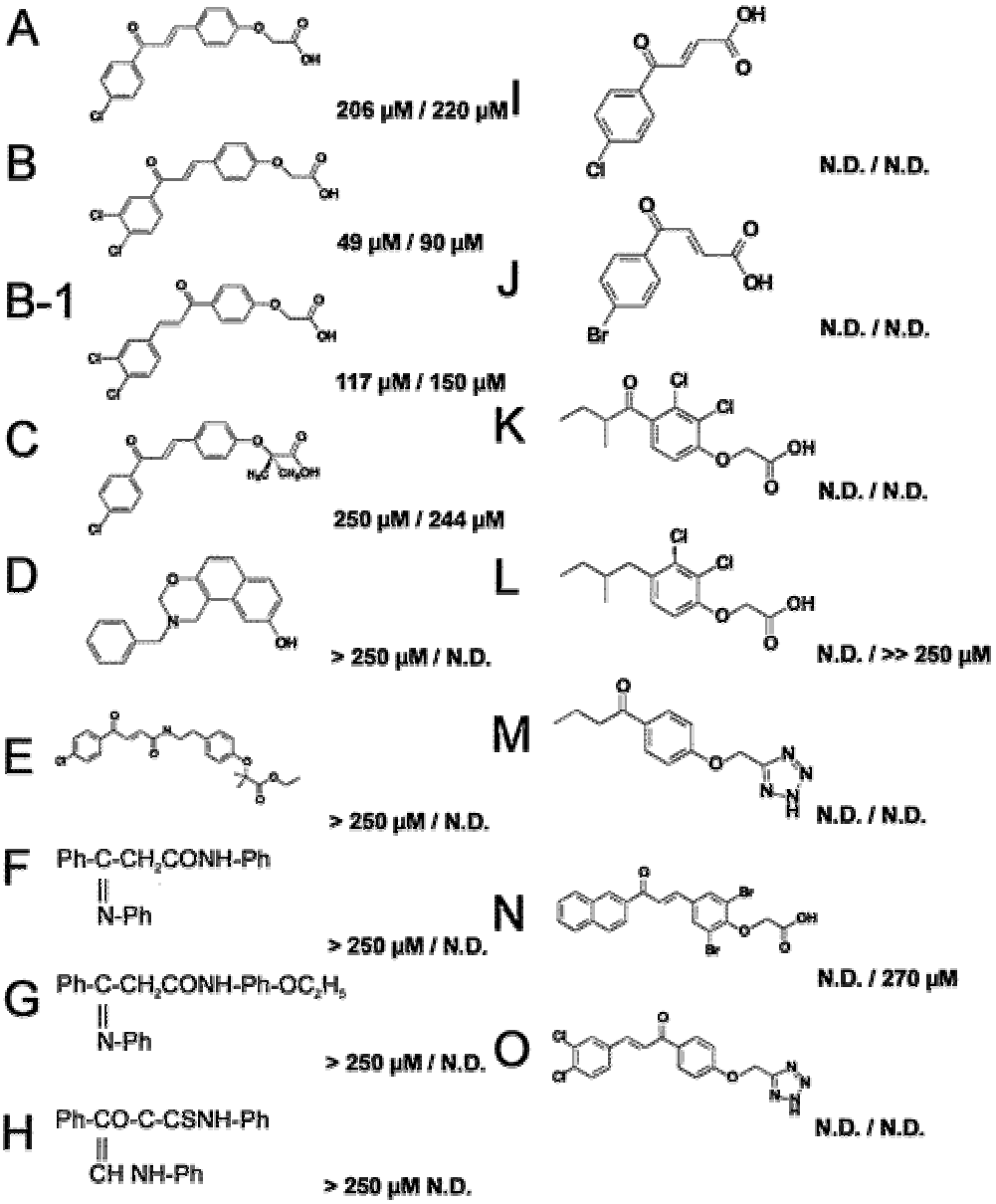


Figure 5.2.4. *A representative collection of basic chalcone skeletons used in our study. Inhibition of MDM2 binding to p53 measured by ELISA (IC_{50} values given on the left side of*

the slash) and by NMR titration experiments (K_D values given on the right side of the slash). Compound D was studied as a negative control. For details refer to text.

The biotinylated optimized p53 peptide, which was coated to a Streptavidin-coated plate, was bound to MDM2 protein. Compounds interfering with the interaction were selected. Since only an 11-mer peptide was used in the ELISA carrying the MDM2 binding site, possible artifacts of secondary or allosteric binding sites are excluded. Because chalcones at certain concentrations induce precipitation/aggregation of MDM2, caution has to be exercised in interpreting the ELISA and EMSA data presented here. However, the ELISA data exhibited a range of IC_{50} values generating a rough structure-activity relationship profile. Compounds B, N, and O denature MDM2. Compound B-1 (Figure 5.2.4) leads to aggregation of the MDM2 but not of the human $p19^{INK4d}$ protein when applied in the same molar excess, as visualized by NMR. Induction of protein aggregation is usually considered as a nonspecific effect of compounds and therefore an indicator of low therapeutic potential. However, aggregation may either arise as a biochemical artifact or as a consequence of a specific interaction. In either case, the substance would inactivate the p53-specific interaction and lead to degradation of cellular MDM2.

5.2.7 Release of p53 Active for DNA Binding by Chalcones

Since the compounds were able to compete with MDM2 for binding to the p53 peptide as shown by ELISA, it was then tested whether they could dissociate preincubated p53/MDM2 complexes and release p53 active for DNA-binding in an electrophoretic gel mobility shift assay (EMSA). Here, full-length, tetrameric p53 protein was used instead of a short peptide. In this setting, the MDM2 protein supershifts p53 bound to its consensus DNA, confirming previously published data (Böttger et al., 1997) (Figure 5.2.1). Since the compounds are dissolved in DMSO, incubation with 5% DMSO was shown not to influence complex formation, as well as a control compound D (Figure 5.2.1). However, the binding of p53 and MDM2 was dissociated by addition of a p53 peptide containing the MDM2 binding site, which thus competes with p53 protein for binding to MDM2 (Figure 5.2.1).

Compounds A and C resolve the p53/MDM2 complex, however, without releasing active p53 (Figure 5.2.1). Thus, the compound interaction seems to additionally influence the p53 protein, which would not be anticipated from the ELISA. Compounds B, N, and O partially remove MDM2 from the complex with p53 by lowering the supershift. The released

p53 still migrates higher than the p53 only control, which could indicate some molecules are MDM2 bound. Despite good IC_{50} values in the ELISA, the effect of the compound seems to be lower in the context of the full-length p53 protein complexed to MDM2. Compound B-1 completely resolves the supershift induced by MDM2 binding and releases active p53.

5.2.8 NMR Spectroscopy

Determination of binding sites of lead chalcone compounds were carried out using ^{15}N-HSQC NMR spectroscopy of the ^{15}N isotopically enriched domain of human MDM2 including residues 1-118. A nearly complete assignment of the backbone $^{1}H^{N}$ and ^{15}N NMR resonances was obtained for the uncomplexed MDM2 (apo-MDM2; Figure 5.2.2). The NMR ^{15}N-{^{1}H} NOE experiment indicated that the folded core of the MDM2 domain in solution extends from T26 to N111 (Figure 5.2.5).

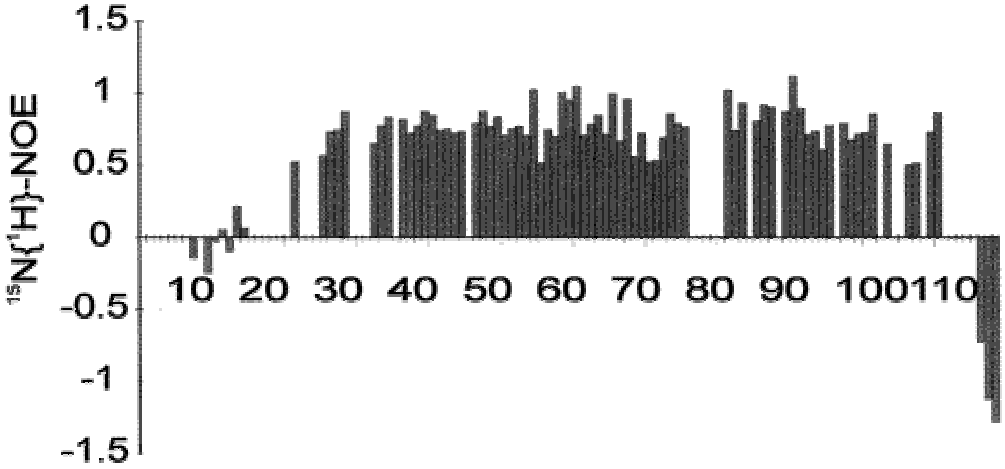


Figure 5.2.5. *^{15}N{^{1}H}-NOE for the backbone amides of human MDM2. Residues for which no results are shown correspond to prolines or to residues where relaxation data could not be extracted.*

This is in good agreement with the crystal structures of N-terminal domains of human and Xenopus MDM2 in complex with a transactivation domain peptide of p53, where the MDM2 structure was also defined from T26 to V109 (Kussie et al., 1996). The p53 peptide, comprising the residues 15 to 29, binds to an elongated hydrophobic cleft of the MDM2 domain. The interaction is primarily hydrophobic in character; only two hydrogen bonds are found between MDM2 and the p53 peptide. The hydrophobic surfaces of MDM2 and p53 are sterically complementary at the interface. The binding surface of p53 is dominated by a triad

of p53 amino acids (F19, W23, and L26) that bind along the MDM2 cleft and define the corresponding phenylalanine, tryptophan, and leucine subpockets for the p53/MDM2 interaction (Kussie et al., 1996) (Figure 5.2.6). In this classification, the leucine pocket is defined by Y100, T101, and V53, the tryptophan pocket is defined by S92, V93, L54, G58, Y60, V93, and F91, the phenylalanine pocket is defined by R65, Y67, E69, H73, I74, V75, M62, and V93 (Kussie et al., 1996).

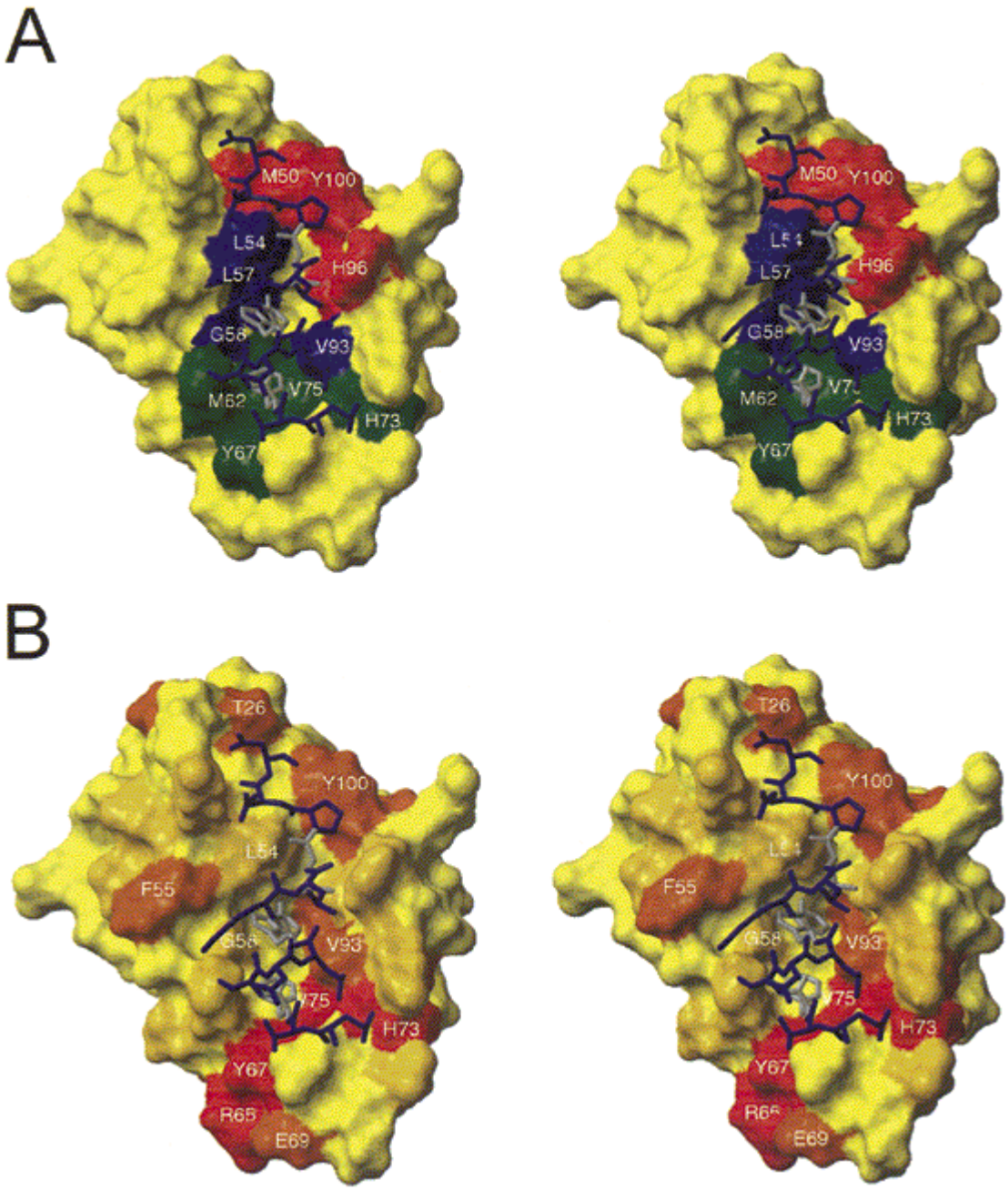


Figure 5.2.6. (A) *Contact surface of human MDM2 (residues 25-109) generated with MOLMOL from the 1YCR data set (Kussie et al., 1996). The atom radius was set to the van*

der Waals value and the solvent radius to 1.4 Å. The p53 peptide is superimposed in blue sticks as a reference with side chain residues of F19, W23, and L26 colored in light blue. Residues of human MDM2 that constitute the leucin-, tryptophan-, and the phenylalanine-binding are colored in red, blue, and green, respectively. (B) *Contact surface of human MDM2 as in panel A. Residues that show significant induced NMR chemical shifts upon complexation with the p53 peptide are highlighted. These residues are shown in yellow, orange, light red, and dark red for observed vectorial shifts smaller than 0.08, 0.08-0.12, and 0.12-0.2 ppm and greater than 0.2 ppm, respectively. The p53 peptide is superimposed in blue sticks as a reference with side chain residues of F19, W23, and L26 colored in light blue.*

As a control experiment using a known stable MDM2/inhibitor complex, MDM2 was titrated with the p53 peptide comprising residues E17 to N29 (Figure 5.2.3, panel A, and Figure 5.2.6). NMR spectra showed that the p53 peptide/MDM2 complex was long-lived on the NMR chemical shift time scale (Wüthrich, 1986; see also Materials and Methods). This is in agreement with the ELISA data that showed an apparent K_D of 0.6 µM (Kussie et al., 1996). As can be seen in Figure 5.2.3, panel A, and Figure 5.2.6, panel B, almost all amino acids of the free MDM2 exhibit changes in chemical shifts upon complexation with the p53 peptide. The analysis of ligand-induced $^{1}H^{N}$ and the ^{15}N shifts was performed by applying the equation of Pythagoras to weighted chemical shifts, which is in concordance with the recent literature (Pellecchia et al., 1999). The largest shifts lined the three binding subpockets of p53 on MDM2 (Figure 5.2.3, panel A, and Figure 5.2.6, panel B). The full set of MDM2/p53 interface residues comprises M50, L54, L57, G58, I61, M62, Y67, H73, V75, F91, V93, H96, I99, and Y100 of MDM2 (Kussie et al., 1996). Additionally, significant shifts are observed for β-strand residues T26, L27, V28, R29, L107, and V108 and for residues L34, L37, and K64 (Figure 5.2.3, panel A, and Figure 5.2.6, panel B). Shifts observed for amides outside the binding regions may be caused by secondary effects, such as allostery or change in mobility upon binding, and do not necessarily indicate direct binding of the p53 peptide to MDM2. Such possible secondary effects (e.g., residues L34, L37, and K64) must be considered when analyzing ligand binding to allosteric proteins.

All K_D values determined by NMR spectroscopy fully agree with the affinities measured by the ELISA binding assay (Figure 5.2.4). Compound A, with an ELISA IC_{50} value of 206 µM, shows the strongest shifts at the peptide groups of E52, V53, L54, F55, Y56, L57, G58, Y60, I61, and H73 (Figure 5.2.4, Figure 5.2.3, panel B, and Figure 5.2.7, panel A).

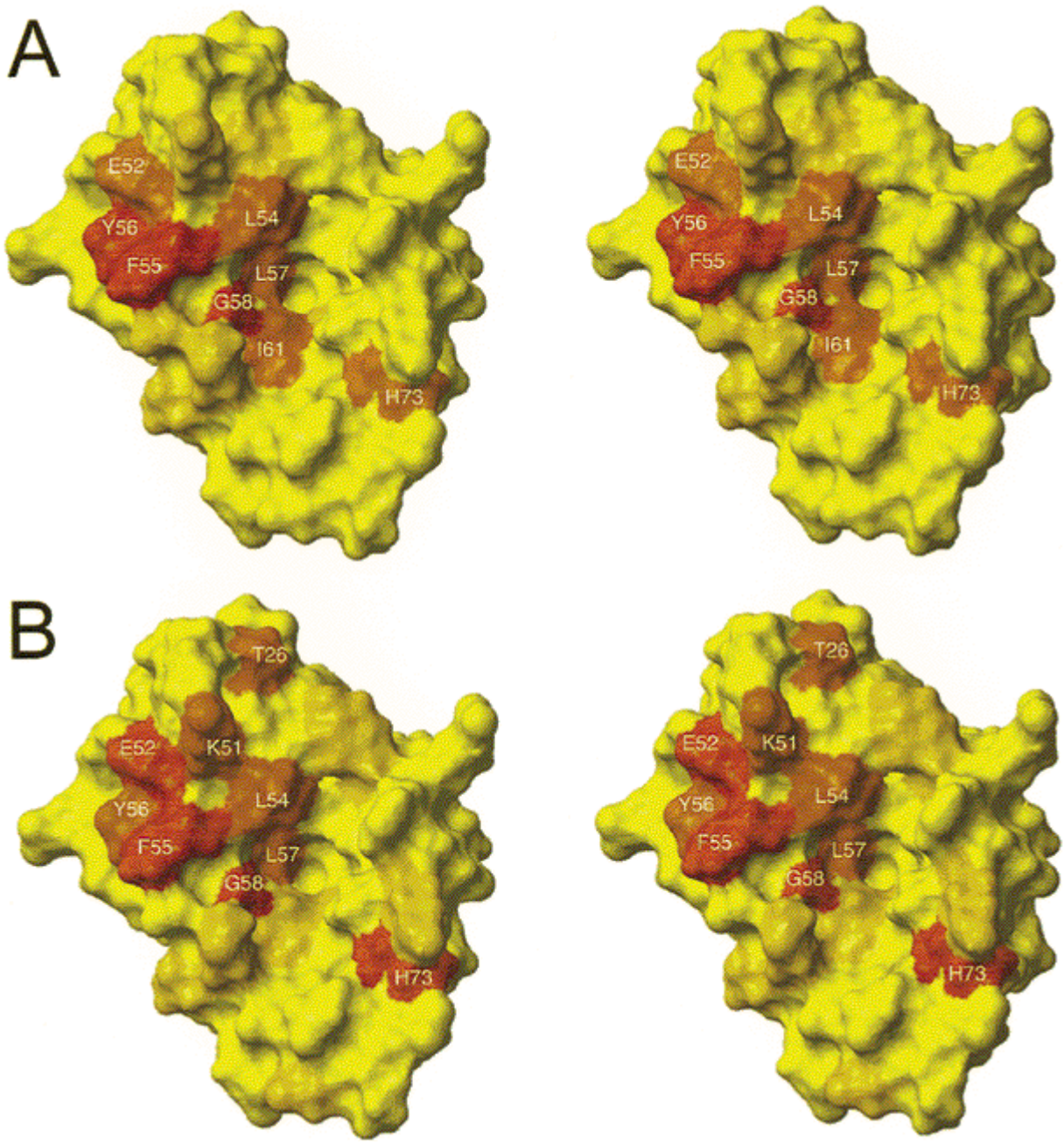


Figure 5.2.7. *Residues that show significant induced NMR chemical shifts upon complexation with chalcone derivatives. These residues are shown in yellow, orange, light red, and dark red for observed vectorial shifts smaller than 0.08, 0.08-0.12, and 0.12-0.2 ppm, and greater than 0.2 ppm, respectively. Contact surface of human MDM2 (residues 25-109) generated with MOLMOL from the 1YCR data set (Kussie et al., 1996). The atom radius was set to the van der Waals value and the solvent radius was set to 1.4 Å. No shift perturbations greater than 0.08 ppm were observed for residues located on the backside of MDM2 for compounds in panels A and B.* (A) *Chalcone A;* (B) *chalcone C.*

Except for H73, all of these are found on the α-helix comprising residues M50-R65; we attribute the H73 shift to secondary or allosteric effects. The shift pattern is consistent with binding in the tryptophan pocket of MDM2. Compounds B and B-1 yielded similar chemical shift patterns as compared to compound A (Figure 5.2.3, panels B, C, and D). The shifts observed for compounds B and B-1 cannot reliably be used to localize the inhibitor interaction site because these inhibitors induce precipitating MDM2/MDM2 interactions that also contribute to the chemical shift pattern. The same is true for compounds N and O.

Chalcone C differs from A by the addition of two methyl groups near the acid terminus, an alteration that insignificantly affects the IC_{50} value (250 μM). The overall NMR shift perturbation pattern is similar to that observed for chalcone A (Figures 5.2.1 and 5.2.3, Figure 5.2.3, panel E, and Figure 5.2.7, panel B). The detailed shift perturbation pattern, however, is changed by the dimethyl substitution: the perturbations observed for T26, K51, and E52 are new or greater, while the perturbations at Y56 and I61 caused by compound C are weakened (Figure 5.2.3, panels B and E, and Figure 5.2.7, panels A and B).

Hypothetical models for the binding modes may be generated using these data (Figures 5.2.7 and 5.2.8). First, a survey of chalcones from the Cambridge Database confirms the overall rigidity and planarity of the extended π-system. Thus, with the assumption described above that the monosubstituted phenyl group binds in the tryptophan pocket, a rotation of the rigid chalcone about the monochlorophenyl group would displace the perturbations from the "lower" region of helix M50-R65 toward the N-terminus to the "upper" region of the helix of the tryptophan subsite. This reflects the perturbation patterns of compound A (including I61 and Y56) and C (T26, K51, E52). Chalcones A and C, docked into the tryptophan subsite, are oriented with acid groups extended toward the solution; the chalcone carbonyl group is also solvent-exposed (Figure 5.2.8). The second phenyl group is also relatively solvent-exposed but encounters the similarly exposed F55 of MDM2 to join a cluster of aromats that further includes Y56. In addition, the acid group can be placed near the base of K51, which is found in a salt bridge interaction with E25 in the crystal structure (Kussie et al., 1996). An intriguing hypothetical possibility is that a salt bridge is formed between K51 and the acid of compound C, with the two methyl groups in a hydrophobic interaction with the aliphatic portion of the lysine side chain. This would break the salt bridge with E25; a conformational change here could cause the amide shift perturbation at T26 as the amide proton is oriented to the same side of the β-sheet T26-P30. Without the two methyl groups of compound C to contribute to K51 binding and compete with E25 for salt bridge formation, compound A would be free to optimize the aromatic group interactions with F55,

Y56, and the tryptophan pocket, leading to a binding conformation in the region and the greater number of perturbations observed for this compound (Figure 5.2.8).

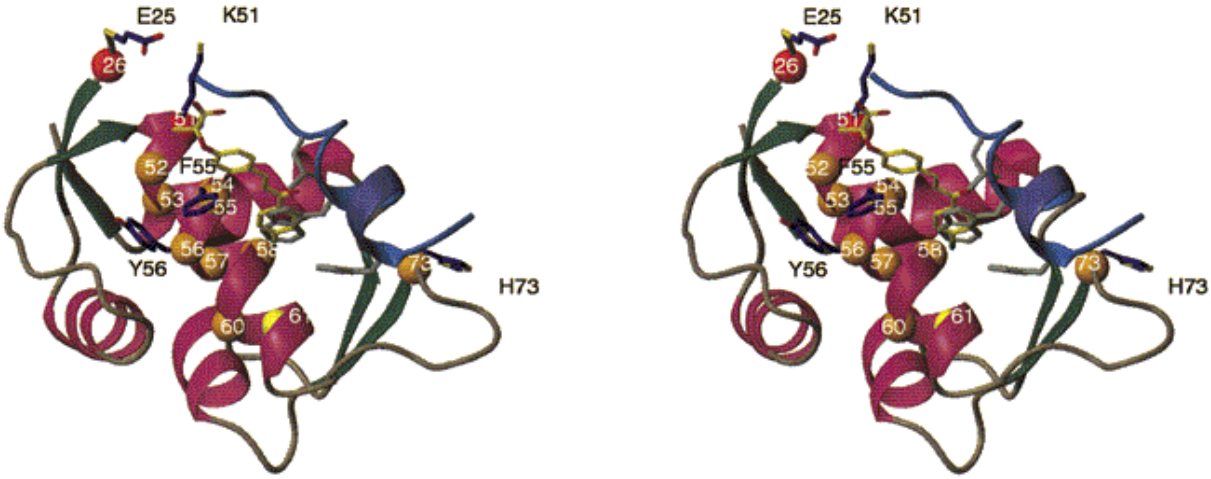


Figure 5.2.8. *Model of human MDM2 in complex with chalcone C (shown in yellow sticks) superimposed with the p53 peptide (shown in blue). The colored spheres indicate residues that showed significant induced chemical shifts upon complexation with the chalcone. For details refer to text.*

Other non-main chain amide (Gln, Asn) shifts are observable in the spectra and can in principle contribute additional information. One such side chain is Q72 that is bound to the p53 peptide in the crystal structure (Kussie et al., 1996). If the outlying amide shift at H73 observed in our experiments is caused by direct ligand binding interactions, the amide of the adjacent side chain Q72 might shift as well. However, for all derivatives of chalcones used in this study, we did not observe any prominent shifts for the side chain protons H_ε of Q72. This is further corroborative evidence for the binding site of A and C in the tryptophan pocket and distant from Q72-H73. Therefore, we conclude that the shifts observed for H73 are secondary and may be caused by changes in the protonation state of the solvent-exposed imidazole ring as the pH of 7.4 the sample was close to the pK_s of the histidine side chain. Another hypothetical explanation for the shifts of H73 observed upon binding of chalcone derivatives is sensitivity to χ-rotamer transitions.

In conclusion, we have shown that chalcone derivatives bind to the tryptophan pocket of the p53 binding site of MDM2 and are able to dissociate the p53/MDM2 complexes. Therefore chalcones, as antagonists of the p53/MDM2 interaction, offer the starting point for structure-based drug design for cancer therapeutics in strategies that abolish constitutive

inhibition of p53 in tumors with elevated levels of MDM2 or, more generally, in strategies that enhance p53 activity.

5.3 Structure of IGF-I and IGFBP-5 Fragment Complex

5.3.1 Protein Expression, Refolding and Purification

Mini-IGFBP-5 (amino acids 40-92 of human IGFBP-5) was expressed using the construct described in Kalus et al. (1998). The mentioned IGFBP-5 fragment was cloned into the *Bam*HI and *Pst*I restriction sites of pQE30 vector (Qiagen, Hiled, FRG) in frame to a His-tag (Kalus et al., 1998). The expression and purification protocol was optimized.

The plasmid was transformed into the BL21 *E.coli* electrocompetent cells. The expression was performed exactly like described in section 4.2.2. The cells were grown until OD_{600} 0.8 was reached, then induced with 1mM IPTG (endconcentration), and incubated for 3 more hours with vigorous shaking (150rpm) at 37°C. Followed the centrifugation for 30minutes at 6000G, and the cell pellet was frozen at –20°C. The pellet derived from 1liter bacterial culture was left overnight in shaker (280rpm) for solubilization in 30ml buffer A/BP5.

Buffer A/BP5

6M guanidine hydrochloride
100mM $NaH_2PO_4*H_2O$
10mM tris, pH 8.0
10mM β-ME

The cell suspension was sonicated (Branson, USA) 2x4 minutes using macrotip, output control 7, 50%. Resulted pellet was centrifuged at 60000G for 1h at 20°C. Supernatant was added to 5ml NiNTA slurry (Qiagen, FRG) equilibrated previously in buffer A/BP5, and shaked gently (130rpm) for 1h at RT. The mixture was then loaded onto an empty column, washed with buffer A/BP5, and subsequently with buffer B/BP5 (protocol like buffer A/BP5 but pH 6.0). The protein was eluted with a gradient to buffer C/BP5.

Buffer C/BP5

6M guanidine hydrochloride
100mM sodium acetate, pH4.5 with acetic acid
10mM β-ME

The fractions containing proteins were detected by Bradford method, pooled and dialysed against buffer D/BP5 to remove reducing agents.

Buffer D/BP5

6M guanidine hydrochloride, pH 3.0

The protein was renatured by rapid dilution of its solution to buffer E/BP5 (1:40) in 1ml steps. The refolding mixture was left gently stirred at 4°C for 3 days.

Buffer E/BP5

200mM L-arginine
1mM EDTA
2mM reduced glutathione
2mM oxidized glutathione
100mM tris-HCl, pH 8.4

The refolding mixture was then centrifuged, concentrated and dialysed to PBS buffer with no NaCl. The pellet was removed by centrifugation, and the supernatant was loaded onto cation exchanger (MonoS, Pharmacia, FRG), and eluted with a gradient to PBS with 1M NaCl. The fractions containing IGFBP-5 were identified with tricine SDS PAGE, pooled and loaded onto the gel filtration column (HiLoad Superdex S75, Pharmacia, FRG) in PBS buffer.

IGF-I was obtained from OvoPepi, Australia.

5.3.2 Crystallization, Data Collection and Derivatization

Crystallization was successful with 10% Jeffamine M-600, 0.1 M sodium citrate, 0.01 M ferric chloride, pH 5.6. Within 11 days, crystals appeared at 4 °C, growing to a final size of about 0.3 x 0.2 x 0.2 mm^3. They belong to the cubic space group $P2_13$ and have unit cell dimensions a, b, c = 74.385 Å, with one complex molecule per asymmetric unit. Soaking in precipitation buffer with heavy atom compounds yielded a derivative K_2PtCl_4 (2.7 mM, 3 d) that was interpretable. All diffraction data were collected using a 300 mm MAR Research (Hamburg, Germany) image plate detector mounted on a Rigaku (Tokyo, Japan) RU300 rotating anode X-ray generator with graphite monochromatized CuKα radiation. All image

plate data were processed with MOSFLM (Leslie 1991) and the CCP4 program suite (Collaborative Computational Project, Number 4 1994).

5.3.3 Phase Calculation, Model Building and Refinement

Table 5.3.1 *Statistics from the crystallographic analysis.*

	native	K_2PtCl_4
Resolution (Å)	16.2 – 2.1	18.6 – 2.5
Measurements	45345	32833
Unique measurements	8035	4925
% Complete (last shell/Å)	99.3 (96.9/2.23 – 2.11)	99.9 (95.4/2.64-2.5)
R_{sym} (%) (last shell)	8.2 (44.8)	8.8 (49.5)
$R_{Cullis\text{-}iso}$	-	0.77
P_{iso}	-	1.48
Res. for phase calc. (Å)	-	18.6 – 2.5
Mean FOM	-	0.41

Refinement statistics:	
Resolution range (Å)	16.2 – 2.1
reflections in working set	7522
reflections in test set	501
R_{cryst} (%)	21.8
R_{free} (%)	26.2
Protein atoms (non-H)	765
Solvent atoms (non-H)	47
Average B-factor (Å^2)	38.1
r.m.s. ΔB (2Å cutoff)	3.4
Deviations from ideality (r.m.s.):	
Bond lengths (Å)	0.013
Bond angles (°)	1.7

$$R_{sym} = \frac{\sum \left| I(h)_i - \langle I(h) \rangle \right|}{\sum \langle I(h) \rangle}$$

$R_{Cullis\text{-}iso}$ = r.m.s. lack of closure / r.m.s isomorphous difference

P_{iso} (Phasing power) = $\langle |F_H| \rangle$ / r.m.s. lack of closure for all reflections

Mean FOM, mean figure of merit.

The structure of the IGF-I/mini-IGFBP5 complex was solved by the single isomorphous replacement (s.i.r.) method using the heavy atom derivative described above. Derivative data was analyzed with the native data set, first using isomorphous difference Patterson maps and employing the Patterson vector superposition methods implemented in SHELX-97 (Sheldrick, 1991). The 2 heavy sites locations were confirmed by difference

Fourier methods with appropriate initial single site s.i.r. phases using CCP4 programs. The refinement of heavy atom parameters and calculation of s.i.r. phases were done with SHARP (La Fortelle et al., 1997). The final parameters are given in Table 5.3.1. The initial s.i.r. phases were improved with SOLOMON (Abrahams & Leslie, 1996) using an solvent fraction of 45%, resulting in a 2.1 Å electron density map that was of such high quality as to enable automated structure building with ARP (Lamzin & Wilson, 1993). All further model building was carried out with the program O (Jones et al., 1991). Refinement was performed by conjugate gradient and simulated annealing protocols as implemented in CNS 1.0 (Brünger et al., 1998). All protocols included refinement of individual isotropic B-factors and using the amplitude based maximum likelihood target function. The R-factor dropped to 21.8 % (R_{free}= 26.2 %, resolution range 16.2 – 2.1 Å) for the final model including 47 water molecules. The water model was calculated using ARP and verified by visual inspection. The final refinement statistics are shown in Table 5.3.1. Coordinates have been deposited in the Protein Data Bank (accession code 1H59).

5.3.4 The IGF-I/mini-IGFBP-5 Complex

Formation of the IGF-I/mini-IGFBP-5 complex buries a binding surface totalling about 550 Å^2. Of the eleven IGFBP-5 residues within 4 Å of IGF, six are hydrophobic, three of which are surface-exposed leucines and valines and are of primary importance for hydrophobic interaction to IGFs (Figures 5.3.1, 5.3.2 and 5.3.3A).

A

Chain B:	**$G_1PETLCGAEL_{10}VDALQF_{16}VCGD_{20}RGFY_{24}FNKPT_{29}$**	**IGF-I**
	$AYRP_4SETLCGGEL_{13}VDTLQF_{19}VCGD_{23}RGFY_{27}FSRPA_{32}$	**IGF-II**
	$FV_2NQHLCGSHL_{11}VEALYL_{17}VCGE_{21}RGFF_{25}YTPK_{29}$	**SCI**
Chain C:	**$G_{30}YGSSSRRAPQ_{40}T$**	**IGF-I**
	S_{33}--RVSRRSR_{40}	**IGF-II**
Chain A:	**$G_{42}IVDECCFR_{50}SCDLR_{55}RLEMY_{60}CA_{62}$**	**IGF-I**
	$G_{41}IVEECCFR_{49}SCDLA_{54}LLETY_{59}CA_{61}$	**IGF-II**
	$G_{A1}IVEQCCTS_{A9}ICSLY_{14}QLENY_{19}CN_{21}$	**SCI**
Chain D:	**$P_{63}LKPAKSA_{70}$**	**IGF-I**
	T_{62}--PAKSE_{67}	**IGF-II**

B

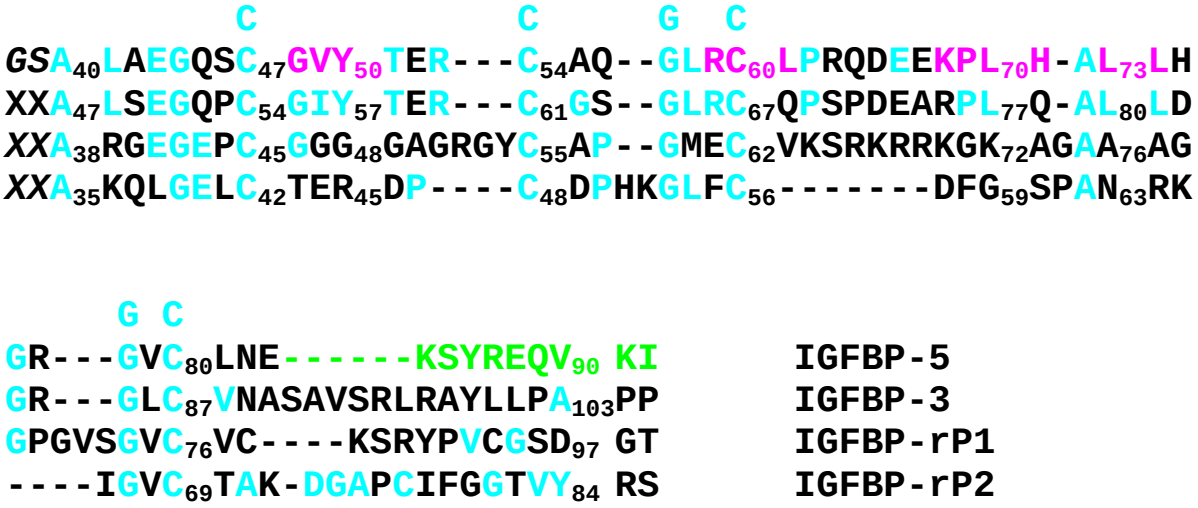


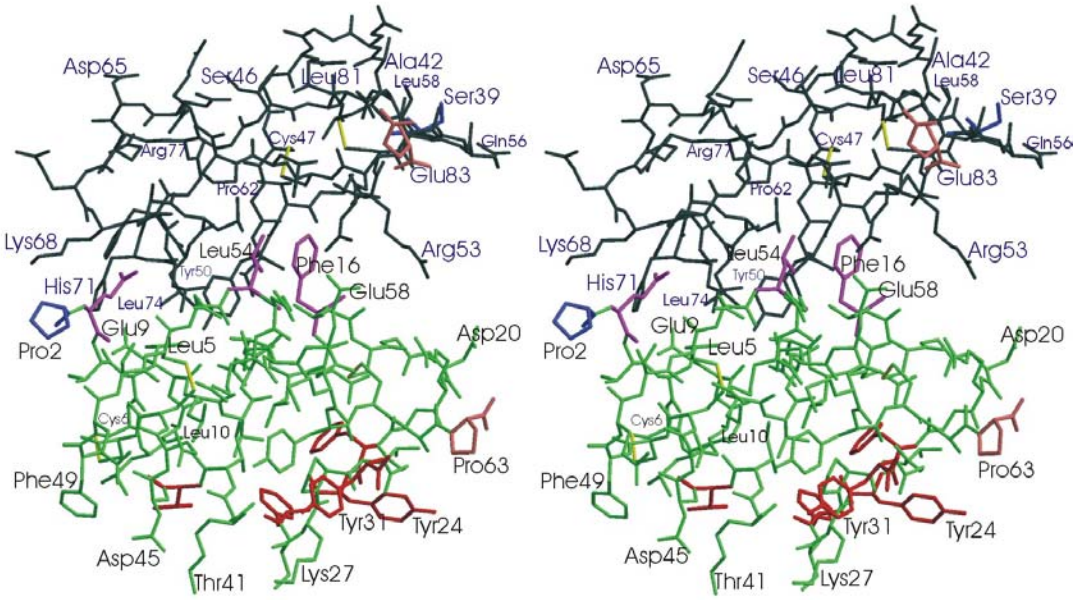
Figure 5.3.1. *Sequence and structure alignment* (A) *of IGFs and single-chain insulin (SCI). Residues that make contacts with mini-IGFBP-5 within 4 Å are highlighted in magenta; residues responsible for binding to IGF-1R in red, residues in green showed no electron density.* (B) *of mini-IGFBP-5 with the corresponding N-terminal domains of IGFBP-3, IGFBP-rP1 and IGFBP-rP2; consensus amino acid residues are shown above the sequences; conserved residues are indicated by blue letters. Residues that interact with IGF-I (within 4 Å) are highlighted in magenta. The mini-IGFBP-5 construct had additional Gly and Ser residues from the cloning vector at the N-terminus; residues in green showed no electron density. Residues shown in parentheses have no structural homology and were aligned based only on amino acid similarities.*

On the IGF side, four of the eleven residues within 4 Å of mini-IBFBP-5 are hydrophobic (Figures 5.3.2B and 5.3.3A).

The principal IGF-I mini-IGFBP-5 interaction is a hydrophobic sandwich that consists of interlaced protruding side chains of IGF-I and solvent exposed hydrophobic side chains of the mini-IGFBP-5 (Figure 5.3.2A). The side-chains of IGF-I Phe 16, Leu 54 and also Glu 3, are inserted deep into a cleft on the mini-IGFBP-5 (Figure 5.3.3A). This cleft is formed by side chains of Arg 53, Arg 59 on the solvent exposed side of the molecule and by Val 49, Leu 70, Leu 74 on the opposite inner side, with a base formed by residues Cys 60 and Leu 61. Phe 16 makes direct contacts with the backbone and side chain of Val 49, and with Cys 60 of mini-IGFBP-5 (Figure 5.3.3). The hydrophobic cluster is closed on the solvent side by side chains of Glu 3 and Glu 9 of IGF-I and His 71 and Tyr 50 of mini-IGFBP-5. These residues form a

network of hydrogen bonds; in addition Arg 59 of mini-IGFBP-5 makes hydrogen bonds with Glu 58 (Figure 5.3.3B).

A



B

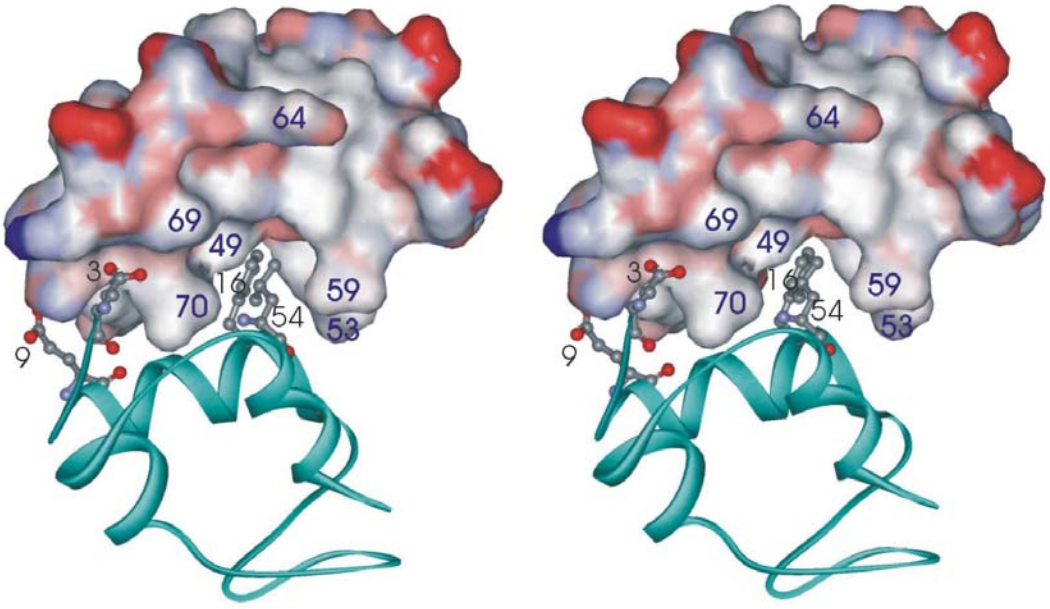


Figure 5.3.2. *The overall structure of the IGF-I (green) mini-IGFBP5 (black) complex.* (A) *A heavy atom plot. Residues shown in magenta constitute the primary binding sites for interaction with mini-IGFBP-5. Residues in red are determinants for binding to IGF-1R. The first N- and last C-terminal residues are shown in brown and blue, respectively.* (B) *Interface of the IGF-mini-IGFBP-5 complex interactions. Mini-IGFBP-5 is shown as a surface plot (residues in red, negatively charged; blue, positive; white, neutral), IGF is shown in blue. Side chains of the primary binding residues of IGF for mini-IGFBP-5 are shown.*

Arg 53 and Arg 59 of mini-IGFBP-5 isolate the hydrophobic sandwich from the solvent close to the C-terminus. In the full length IGFBP-5, the segment corresponding to the C-terminus of mini-IGFBP-5 is followed by nine hydrophilic residues and then by at least 30 residues of mixed types. Thus we can postulate that the conformations seen in the structure of the complex near the C-terminus of mini-IGFBP-5 are likely to be preserved in the complex of IGF-I with the full length-IGFBP-5. The mini-IGFBP-5 domain begins at residue 40 of full length IGFBP-5. Our previous NMR study of binding of the N-terminal domain of IGFBP-5 (from residues 1–102) showed unequivocally that this 39-residue segment did not interact with IGFs and that the first 39 residues of IGFBP-5 have no influence on the structure of the following mini-IGBFP-5 domain (Kalus et al., 1998).

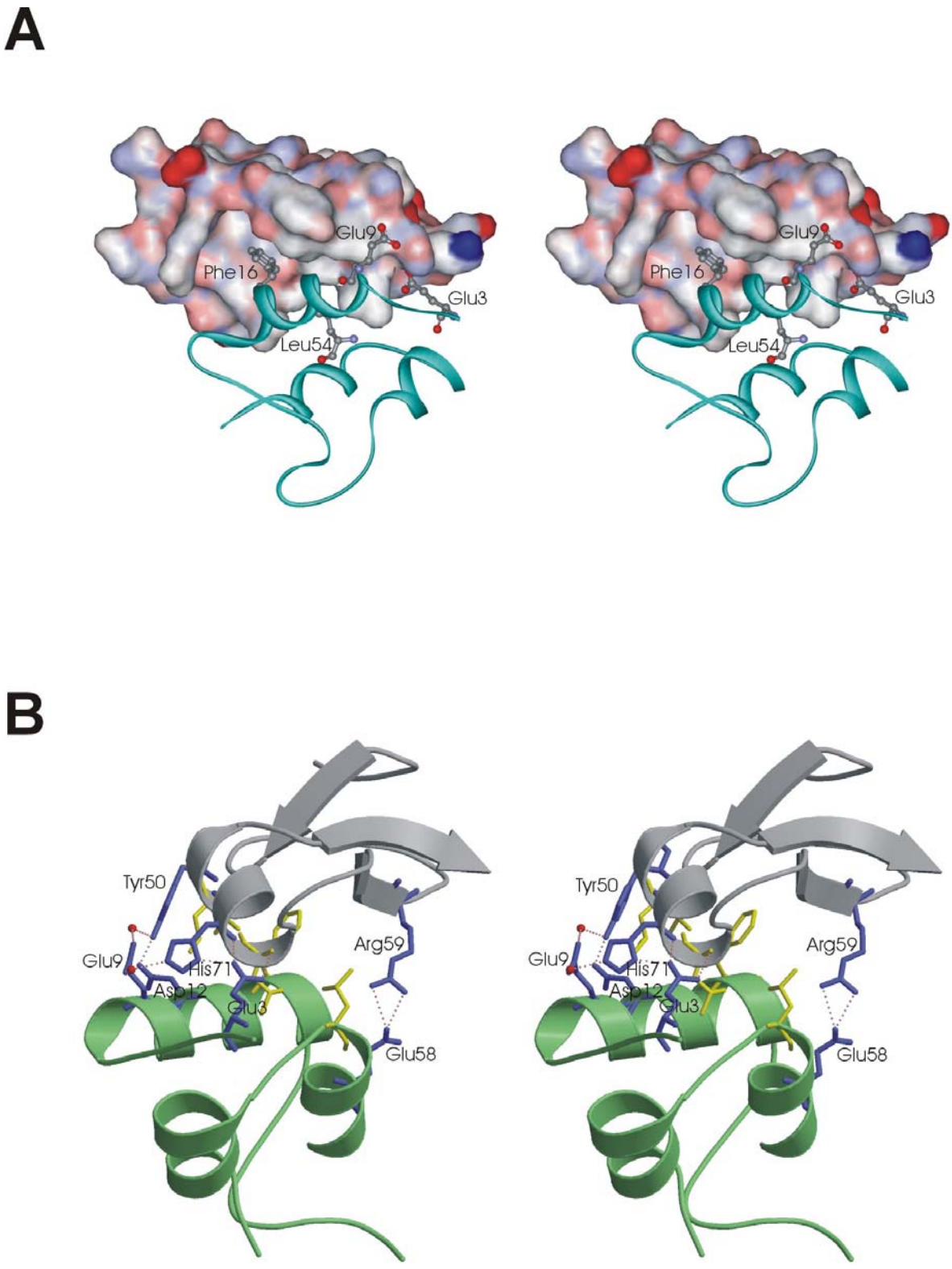


Figure 5.3.3. *Interface of the IGF-mini-IGFBP-5 complex interactions.* (A) *Mini-IGFBP-5 is shown as a surface plot (residues in red, negatively charged; blue, positive; white, neutral), IGF is shown in blue. Side chains of the primary binding residues of IGF for mini-IGFBP-5 are shown. The molecule is rotated 180^{o} around the vertical axis compared to Figure 5.3.2B.* (B) *Ribbon plot of IGF (green) mini-IGFBP-5 (gray) with interface residues that form*

hydrogen bonds are highlighted (blue). The interface hydrophobic residues are shown in yellow.

Mutagenesis studies for IGFs indicated that IGF residues Glu 3, Thr 4, Gln 15 and Phe 16 of IGF-I and Glu 6, Phe 48, Arg 49 and Ser 50 in IGF-II are important for binding to IGFBPs (Baxter et al., 1992). Baxter et al. (1992) suggested that the IGF-I Glu 3, Thr 4, Gln 15 and Phe 16 are crucial for interaction with IGFBP-3, whereas residues Phe 49, Arg 50 and Ser 51 are of secondary importance. It also was suggested that Phe-26 of IGF-II plays a role in changing the local structures of IGFs but does not directly bind to IGFBPs (Terasawa et al., 1994). Not all of these residues make direct contacts (within 4 Å) with mini-IGFBP-5: of the residues identified by mutagenesis, Gln 15 neighbours the important Phe 16, the IGF-I residues Phe 49, Arg 50 and Ser 51 (equivalent to IGF-II 48, 49, 50) are within three residues from the interface, and Phe 23 (IGF-II Phe 26) is far from the complex contact.
Our previous NMR study showed that the hydrophobic residues Val 49, Leu 70 and Leu-73 of IGFBP-5 are crucial for binding to IGFs, which is fully in agreement with the current structure. Since these residues are highly conserved among all IGFBPs we expect that these hydrophobic interactions dominate the IGF binding properties of all IGFBPs and also IGFBP-rPs. For IGFBP-rPs, it possible to produce a model of the structure of the N-terminal domains bound to IGFs using the structure of mini-IGFBP-5 as a template (data not shown; c.f. Figure 5.3.1B). In IGFBP-rP1, the crucial Leu 70 of IGFBPs is replaced by Lys 72. In the model of the complex β and γs of Lys 72 make hydrophobic contacts to IGF residues sideway similarly to Leu 70 of IGFBPs. The terminal NH_2s of Lys 72 can insert deep into the pocket of IGF-I.

5.3.5 2.1Å Resolution Atomic Structure of IGF-I

The general fold of the free IGF-I found in the best NMR structure, that of long-[Arg3]IGF-I (Laajoki et al., 2000), is preserved in the complex, but the average root mean square deviations (r.m.s.d.) between the NMR and the X-ray structures for well-defined parts of the NMR structures (residues 3-25 and 41-63) is high with 3.7±1.6 Å for α-carbons. Regrettably, the coordinates of the best quality NMR structure of IGFs, that of IGF-II, are not available (Terasawa et al., 1994). For these structures, the ensemble of the structures seems to be highly defined for most of the residues. However, large variabilities in the structures were seen for residues 1-6, the C-terminal residues 62-67, and most importantly, for the chain C residues 31-40 that form a peripheral loop. This is interesting because most of the C chain and

the C-terminus in our IGF-I structure (e.g. residues 32-40 and 64-70) showed no electron density (Figure 5.3.1A and 5.3.2B). The NMR and the present X-ray data indicate therefore an increased motional flexibility in these regions of the IGF molecules. The N-terminal residues of IGF-I are well defined in the X-ray structure of the complex and unstructured in the NMR structures. The N-terminus of IGFs includes several key residues responsible for the interaction to IGFBPs and therefore the conformations of these residues most probably become fixed only upon complex formation.
The model structure constructed by Blundell et al. (1978) is closest to the present X-ray structure. R.m.s.d. values for residues between 3-25 and 41-63 are 1.07 Å for α-carbons and 2.2 Å for heavy atoms. The side chain conformation of Phe-16, the residue responsible for primary interactions with IGFBP-5, is similar in both structures, however, the conformations of Glu 3 and Leu 54 differ, although the χ^1 rotamers are similar in both structures.

5.3.6 Comparison Between Complexed and Free mini-IGFBP-5

The fold of the uncomplexed mini-IGFBP-5 determined by NMR (Kalus et al., 1998) is preserved in the complex. A solvent exposed loop between Pro 62 and Pro 69 was the least precisely defined segment of the structure, and five C-terminal residues of mini-IGFBP-5 were unstructured. ^{15}N relaxation measurements indicated that the backbone of the variable loop 62-69 does not exhibit any fast picosecond time scale motions; instead, the loop residues in the IGF-free mini-IGBP-5 show contributions from slower exchange processes with millisecond range (data not shown). IGF complex formation, however, rigidifies of this loop. In the crystal structure, the loop adopts one of the many conformations that were possible for the free mini-IGFBP-5.

5.3.7 Implications for IGF Binding to Its Receptor (IGF-IR)

The IGF-I receptor (IGF-1R) is a transmembrane heterotetrameric protein complex that has approximately 60% sequence homology to the insulin receptor (IR). IGF-1R binds also IGF-II and insulin with 2- to 15- and 1000-fold lower affinity, respectively (Khandwala et al., 2000). There is also an IGF-II specific receptor: the IGF-II/mannose 6-phosphate receptor, a monomeric receptor that binds IGF-II with a 500-1000-fold increased affinity over IGF-I but does not bind insulin. Most of the actions of IGF-II are however believed to be mediated via the IGF type 1 receptor (Khandwala et al., 2000). Since the ligands IGF-I, IGF-

II and insulin share a common architecture and cross-react with IGF-1R and IR, it is thought that they bind to these receptors in a structurally equivalent fashion (Torres et al., 1995; Gill et al., 1996).

Extensive site directed mutagenesis studies of mapping binding sites of IGFs and insulin for IGF-1R and IR showed that the major determinants of binding are located in the N-terminal region of the A-chain and the C-terminal strand of the B-Chain (Murray-Rust et al., 1992). In IGF-I, the three aromatic residues Phe 23, Tyr 24, and Phe 25 are known to be crucial for receptor binding (Cascieri et al., 1988) and also the A-chain Val 44 is important for binding (Figure 5.3.2). Bayne et al. (1990) have demonstrated that the IGF-1 receptor recognizes in addition Tyr 31 and Tyr 60; in fact, all three tyrosines (24, 31 and 60) are protected from iodination when bound to IGF-1R, indicating that these residues are part of or are near to the binding site. The C region of IGF-I seems to be important in maintaining high affinity binding to the type 1 IGF receptor, since the replacement of the C region of IGF-I with a four glycine span such as in [1-27,Gly4,38-62]hIGF-I results in a 30-fold loss of affinity for IGF-1R. More recently it was shown that binding to the IGF receptor is lost in a "mini" deletion construct of IGF-I in which Pro 28 and Gly 42 are peptide linked. Removal of the D region has little effect on binding to IGF-1R.

Figures 5.3.1A and 5.3.2A show the location of the residues involved in the IGF-1R binding in our IGF-I structure. In most cases these residues correspond to those mapped on the structures of IGFs derived previously from NMR studies (Cooke et al., 1991; Sato et al., 1993; Laajoki et al., 2000). A general trend established from comparing IGFs binding to IGF-1R, IGF-2R, IR, and IGFBPs was that the residues that bind to the type 1 receptor appear to overlap those that bind to the insulin receptor, whereas those that bind to type 2 receptor overlap those that interact with IGF binding proteins.

The most notable features evident from Figure 5.3.2 is that the binding site for IGF-1R consists of a fully solvent exposed hydrophobic patch that is located on the opposite side of IGF to that for the binding to mini-IGFBP-5. This is in contrast to insulin where the binding site for IR is partially occluded by the C-terminus of the B-chain and it is now uniformly accepted that the C-terminus moves away from the surface of the insulin monomer on receptor binding and makes the highly conserved side chains of Ile 2 and Val 3 accessible for binding (Hua et al., 1991).

The current structure supports also an attractive explanation of the results of our studies on inhibition of the IGF binding to the IGF-I receptor by IGFBP-5 and mini-IGFBP-5 and on the influence of the IGFBP-5/IGF complex formation on IGF-mediated stimulation of

the IGF-1R auto-phosphorylation (Kalus et al., 1998). Whereas a complete inhibition of IGF1R-IGF binding was observed as soon as IGFBP-5 was in excess to IGF, a 10^3-fold excess of mini-IGFBP-5 was needed to block IGF-II binding to its receptor. Obviously, the C-terminal domain of IGFBP-5 is essential for effective inhibition of receptor binding of IGF-I. In addition, incomplete inhibition of receptor binding was observed for mini-IGFBP-5 even at the highest concentrations used. From our structure, it appears that IGF-I can still freely bind to the receptor even when complexed to the truncated IGFBP-5 fragment. The lower inhibitory potency of mini-IGFBP-5, compared to the full-length IGFBP-5, would also be decreased by its 10-fold reduction in the binding to IGF-I.

5.3.8 Implication for Therapeutic Modulation of the GH/IGF System for Stroke and Tumorigenesis

The present structure enables *in silico* screens for small IGFBP ligand inhibitors with the potential to release "free" bioactive IGF-I. Displacement of IGF from their binding proteins in brain tissue, for example, should have therapeutic benefits for stroke and other neuro-degenerative diseases. It has recently been demonstrated that a high-affinity IGFBP ligand inhibitor, [Leu24,59,60, Ala31]hIGF-I, that binds to IGFBPs but not to IGF-1R, elicits neuroprotective effects comparable to those produced by the administration of exogenous IGF. In a rat model of focal ischemia, administration of this analog after ischemic insult to the rat brain had potent neuroprotective action comparable to IGF-I (Loddick et al., 1998).

The association of insulin-like growth factors with neoplasia indicates that modulation of the IGFBP environment might be a successful strategy in cancer therapy. Such modulation might be accomplished, for example, through exogenous administration of recombinant protein effective fragments. Additionaly, tumor cell IGFBP production, inhibition or degradation may be controlled by agents such as tamoxifen and ICI 182,780 that modify tumor IGFBP production (Khandwala et al., 2000). The consequent alteration in IGFBP-3 levels appears in certain instances to inhibit IGF-I-stimulated cell proliferation (Khandwala et al., 2000). There is also recent evidence that IGFBP-3 may be a p53-independent effector of apoptosis in breast cancer cells via its modulation of the Bax:Bcl-2 protein ratio (Butt et al., 2000; Wetterau et al., (1999).

Based on the knowledge of the mini-IGFBP-5 structure (Kalus et al., 1998), mutants have been produced with modulated IGF-action and altered cleavage susceptibility for IGFBP-5 protease (Imai et al., 2000). Such mutants may identify roles for IGFBPs that

require IGF-I binding and distinguish them from those that are IGF independent (Imai et al., 2000). In conclusion, the structure of the IGF-I mini-IGFBP-5 complex will advance the development of IGFs with reduced binding affinity for IGFBPs and consequently enhanced activity and of IGFBPs with higher affinity for IGFs and consequent inhibition of IGF signalling. Furthermore, it should contribute to the search for small IGFBP ligand inhibitors, which release IGFs from the inactive complex with IGFBPs.

6 Summary

In this thesis, three human proteins involved in the regulation of cell division and consequently in cancer, E2F-1, MDM2, and IGF-I, were investigated. The proteins create a complex network of interacting factors. E2F-1 is activated by interaction with MDM2. Overexpression of E2F-1 can lead to the p53 dependent apoptosis, which is, in turn, negatively controlled by MDM2-p53 interactions. E2F-1 is indirectly positively regulated by mitogens, example of which is IGF-I. IGF-I gene expression is, however, transactivated by E2F-1. This shows how complicated molecular basis of cancer are and shows that looking for a cure against the disease should be performed on many levels.

E2F transcription factors regulate expression of a group of cellular genes that control cellular DNA synthesis and proliferation. They include cellular oncogenes, tumor suppressor genes, and other rate-limiting regulators of DNA synthesis and cell cycle progression. E2F-1 plays also an essential role in cell differentiation, and forced induction of E2F activity leads to apoptosis. A functional E2F transcription factor consists of a heterodimer containing an E2F and a DP polypeptide. Each of these polypeptides can bind to the E2F-specific DNA sequences and stimulate transcription when overexpressed. The retinoblastoma susceptibility protein (RB) has been reported to alter the functions of E2F by direct binding, which leads to the repression of the transactivation, repression of apoptosis, protection from degradation and altered E2F-DNA binding site specificity. The viral oncoproteins mediate their action through the activation of the E2F by disruption of the RB/E2F complex. Given the functions of E2F-regulated genes, it is reasonable to expect that their altered expression could contribute to the development of cancer. In this thesis preliminary studies were carried out on protein expression and purification of the members of the E2F family. E2F-1 protein constructs in *E. coli* were shown to be well overexpressed but fully insoluble. The trials to optimize solubility of a GST-fused E2F-1 DNA-binding fragment were successful. The fragment of interest appeared, however, to be insoluble after GST cleavage. Attempts to purify and refold the insoluble histidine-tagged full length E2F-1 resulted in the pure and soluble, however unfolded polypeptide as shown by NMR spectroscopy. Additional studies showed that baculovirus expression vectors system could be an other way of producing the biologically active and properly folded protein.

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

Insulin-like growth factors (IGFs) are key regulators of cell proliferation, differentiation and transformation, and are thus pivotal in cancer, especially breast, prostate and colon neoplasms. They are also important in many neurological and bone disorders. Their potent mitogenic and anti-apoptotic actions depend primarily on their availability to bind to the cell surface IGF-I receptor. In circulation and interstitial fluids, IGFs are largely unavailable, as they are tightly associated with IGF-binding proteins (IGFBPs) and are released after IGFBP proteolysis. This thesis presents the 2.1 Å crystal structure of the complex of IGF-I bound to the N-terminal IGF-binding domain of IGFBP-5 (mini-IGFBP-5), a prototype interaction for all N-terminal domains of the IGFBP family. The principal interactions in the complex comprise interlaced hydrophobic side chains that protrude from both IGF-I and the IGFBP-5 fragment and a surrounding network of polar interactions. A solvent-exposed hydrophobic patch is located on the IGF-I pole opposite to the mini-IGFBP-5 binding region and marks the IGF-I receptor binding site. The structure led to the design of small organic compounds that were shown to disrupt the binding of IGFBP-5 to IGF-I, thus, providing starting leads for therapeutics that could influence the regulation of the IGF-IGFBP axis.

7 Zusammenfassung

Im Rahmen dieser Doktorarbeit wurden drei humane Proteine E2F-1, MDM2, IGF-I, die an der Regulierung der Zellteilung und dadurch an der Krebsentstehung beteiligt sind, untersucht. Diese Proteine wechselwirken untereinander in einer komplexen Weise. E2F-1 wird durch das MDM2-Protein aktiviert. Eine Überexpression des E2F-1-Proteins kann zu einer vom p53-Protein abhängigen Apoptose führen. Diese Apoptose wird im Gegenzug durch die MDM2-p53-Wechselwirkung eingeschränkt. Eine Aktivierung von E2F-1 erfolgt indirekt über Mitogene wie zum Beispiel IGF-I. E2F-1 transaktiviert die IGF-I-Genexpression. Diese komplexen Abhängigkeiten demonstrieren die Schwierigkeiten das Verständnis der Entstehung von Krebs auf molekularer Basis zu überwinden. Ein Erkenntniszuwachs wird voraussichtlich nur in vielen kleinen Schritten erfolgen.

Die E2F-Transkriptionsfaktoren regulieren die Expression einer Gruppe von zellularen Genen, die DNA-Synthese und Zellprolifertration kontrollieren. Dazu gehören zellulare Onkoproteine, Tumorsuppressorgene und andere Regulatoren der DNA-Synthese und Progression im Zellzyklus. Der Transkriptionsfaktor E2F-1 spielt außerdem eine wichtige Rolle in der Zelldifferentiation. Eine verstärkte Induktion der E2F-Aktivität führt zur Apoptose. Der biologisch aktive E2F-Transkriptionsfaktor besteht aus einem Heterodimer, das von einem E2F-Molekül und einem DP-Polypeptid gebildet wird. Jedes dieser Polypeptide kann an E2F-spezifische DNA-Sequenzen binden und stimuliert im Falle der Überexpression die Transkription. Das "retinoblastoma susceptibility protein" (RB) kann die Funktionen des E2F durch direkte Bindung verändern. Dies führt zur Unterdrückung der Transaktivierung und der Apoptose. Die Bindung von RB an E2F verhindert die Degradation und verändert die Spezifität der DNA-Bindestelle in E2F. Virale Onkoproteine sind in der Lage den RB/E2F-Komplex zu zerstören, was zur Aktivierung des E2F führt. Wenn man die Funktionen der von E2F regulierten Gene betrachtet, ist es sehr wahrscheinlich, dass deren veränderte Expression eine Rolle bei der Entstehung von Krebs spielt. In dieser Dissertation wurden einleitende Studien zur Expression und Reinigung von Proteinen der E2F-Familie durchgeführt. Verschiedene Konstrukte von E2F-1 wurden in *E. coli* in hoher Ausbeute aber vollkommen unlöslich exprimiert. Versuche zur Erhöhung der Löslichkeit eines mit GST fusionierten und DNA-bindenden Fragments von E2F-1 waren erfolgreich. Nach enzymatischer Entfernung von GST erwies sich das Fragment als unlöslich. Versuche zur Reinigung und Rückfaltung des vollständigen, mit Hexahistidin-Tag fusionierten, E2F-1

führten zu einem sehr reinen und löslichen Polypeptid, das laut NMR-Spektroskopie keine Faltung aufwies. Weiterführende Studien zeigten, dass das Baculovirusexpressionsvektorsystem eine Alternative zur Produktion von biologisch aktivem Protein darstellt.

Das Onkoprotein MDM2 inhibiert das Tumorsuppressorprotein p53 durch Bindung an dessen Transaktivierungsdomäne. Das Protein p53 ist in vielen Tumoren durch Mutationen oder Bindung an Onkoproteine inaktiviert. In manchen Tumoren wie Gewebe-Sarkomas inhibiert die Überexpression von MDM2 aktives p53. Dies bewirkt die Abschaltung des Kontrollpunktes für Schäden am Genom und zur Einleitung der Zellteilung von Zellen mit Gendefekten. Die Zerstörung des MDM2/p53-Komplexes stellt die Aktivität des Tumorsuppressorproteins wieder her, was die therapeutische Bedeutung von MDM2/p53-Bindungsantagonisten unterstreicht. Die Untersuchungen mit Hilfe der NMR-Spektroskopie in der vorliegenden Arbeit zeigten, dass Chalcone (1,3-diphenyl-2-propen-1-one) in einem bestimmten Bereich der p53-Bindetasche des humanen MDM2 binden können. Biochemische Experimente bewiesen die Fähigkeit dieser Verbindungen, p53 aus p53/MDM2-Komplexen und aus an DNA gebundenen p53/MDM2-Komplexen freizusetzen. Diese Ergebnisse bieten die Grundlage für die Entwicklung neuer Wirkstoffe in der Krebstherapie.

"Insulin-like growth factors" (IGFs) sind Schlüsselregulatoren der Proliferation, Differentiation und Transformation von Zellen. Sie spielen damit eine zentrale Rolle bei der Entstehung von Neoplasmen in Brust, Prostata und Dickdarm sowie bei der Entwicklung von Krankheiten am Skelett und neurologischer Funktionsstörungen. Ihre stark mitogene und antiapoptotische Wirkung hängt von ihrer Fähigkeit zur Bindung an IGF-I-Rezeptoren an der Zelloberfläche ab. Im Blutkreislauf und interstitiellen Flüssigkeiten sind IGFs größtenteils an IGF-bindende Proteine (IGFBPs) assoziiert und werden erst nach Proteolyse der IGFBPs freigesetzt. In der Dissertation wurde die Kristallstruktur des Komplexes zwischen IGF-I und der N-terminalen IGF-Bindungsdomäne von IGFBP-5 (mini-IGFBP-5) mit einer Auflösung von 2,1 Å aufgeklärt. Diese Wechselwirkung mit IGF-I ist typisch für alle N-terminalen Domänen der IGFBP-Familie. An den grundlegenden Wechselwirkungen sind ineinander verflochtene hydrophobe Seitenketten, die aus beiden Komplexpartnern herausragen und ein umgebendes Netzwerk aus polaren Interaktionen, beteiligt. Eine zum Solvenz hin exponierte hydrophobe Region befindet sich am zur IGFBP-5-Bindestelle entgegengesetzten Ende und markiert die Bindestelle des IGF-I-Rezeptors. Die Struktur führte bereits zur Entwicklung von

kleinen organischen Molekülen, die in der Lage sind die Assoziation von IGBP-5 an IGF zu unterbinden. Dies eröffnet die Möglichkeit für therapeutische Ansätze zur Regulation des IGF/IGFBP-Systems.

8 Appendix: Abbreviations and Symbols

- Å Ångstrøm (10^{-10}m)
- aa amino acid
- ALS acid labile subunit
- ATP adenosine triphosphate
- 1D one-dimensional
- APS ammonium peroxodisulfate
- BEVS baculovirus expression vectors system
- bHLH basic region and helix-loop-helix region
- bp base pair
- BSA bovine serum albumin
- CAK CDK activating kinase
- CDC cyclin-dependent kinase
- cDNA complimentary DNA
- CDK cyclin-dependent kinase
- chalcone 1,3-diphenyl-2-propen-1-one
- CIP inhibitor of kinase
- COSY correlation spectroscopy
- CSF colony-stimulating factor
- δ chemical shift
- Da Dalton (g mol^{-1})
- DHFR dihydrofolate reductase
- DMSO dimethylsulfoxide
- DNA deoxyribonucleic acid
- DNaseI deoxyribonuclease I
- DP dimerization protein
- DTT Dithiothreitol
- E2F factor interacting with adenovirus E2-promoter
- EDTA ethylenediamine tetraacetic acid
- EGF epidermal growth factor
- ELISA enzyme-linked immunosorbant assay
- EMA E2F-binding site modulating activity

- EMSA electrophoretic gel mobility shift assay
- FID free induction decay
- G gravity (9.81 m s^{-2})
- GH growth hormone
- GSH reduced glutathione
- GSSG oxidized glutathione
- GST glutathione S-transferase
- HA hemagglutinin
- HAT histone acetyltransferase activity
- HDAC histone deacetylase activity
- HPV human papillomavirus
- HSQC heteronuclear single quantum coherence
- Hz Hertz
- IGF insulin-like growth factor
- IGFBP IGF binding protein
- IGFBPrP IGFBP related protein
- IGF-IR IGF receptor type I
- INK4 inhibitor of CDK4 and CDK6
- IPTG isopropyl-β-thiogalactopyranoside
- IRS insulin receptor substrates
- KIP inhibitor of kinase
- LB Luria-Broth medium
- M mol l^{-1}
- MAD multiwavelength anomalous diffraction
- MAP mitogen-activated protein kinase
- MCM minichromosome maintenance protein
- MDM2 murine double minute clone 2
- MIR multiple isomorphous replacement
- MM minimal medium
- MW molecular weight
- NiNTA nickel-nitrilotriacetic acid
- NLS nuclear localization signal
- NMR nuclear magnetic resonance

- NOE nuclear Overhauser effect
- NOESY nuclear Overhauser enhancement spectroscopy
- OD optical density
- P3K phosphatidylinositol 3-kinase
- PAGE polyacrylamide gel electrophoresis
- PBS phosphate-buffered saline
- PCNA proliferating cell nuclear antigen
- POL DNA polymerase α
- ppm parts per million
- RB retinoblastoma susceptibility protein
- RF RING finger motif, or radio frequency
- RMSD root mean square deviation
- RNaseA ribonuclease A
- SAR structure-activity relationship
- SDS sodium dodecyl sulfate
- SV 40 simian virus 40
- TBP TATA-binding protein
- TEMED N,N,N',N'-tetramethylethylendiamine
- TFIIE transcription factor IIE
- TK thymidine kinase
- TOCSY total correlation spectroscopy
- TS thymidylate synthase
- ZF zinc finger motif

Amino acids and nucleotides are abbreviated according to either one or three letter IUPAC code.

9 References

Abrahams, J.P. & Leslie, A.G.W. (1996) Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta. Cryst*. **D52**, 30-42

Arunkumar, A.I., Kumar, T.K., Jayaraman, G., Samuel, D. & Yu. C. (1997) Induction of helical conformation in all beta-sheet proteins by trifluoroethanol. *J. Biomol. Struct. Dyn.* **14**, 381-385

Bagchi, S., Weinmann, R. & Raychaudhuri, P. (1991) The Retinoblastoma Protein Copurifies with E2F-I, an E1A-regulated Inhibitor of the Transcription Factor E2F. *Cell* **65**, 1063-1072

Bandara, L.R. & La Thangue, N.B. (1991) Adenovirus E1a prevents the retinoblastoma gene product from complexing with a cellular transcription factor. *Nature* **351**, 494-497

Baumgartner, R., Fernandez-Catalan, C., Winoto, A., Huber, R., Engh, R.A. & Holak, T.A. (1998) Structure of human cyclin-dependent kinase inhibitor p19INK4d: comparison to known ankyrin-repeat-containing structures and implications for the dysfunction of tumor suppressor p16INK4a. *Structure* **6**, 1279-1290

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

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

Blundell, T.L., Berdarkar, S., Rinderknecht, E. & Humblel, R.E. (1978) Insulin-like growth factor: a model for tertiary structure accounting for immunoreactivity and receptor binding. *Proc. Natl. Acad. Sci. USA* **75**, 180-184

Bois, F., Boumendjel, A., Mariotte, A.-M., Conseil, G. & Di Pietro, A. (1999) Synthesis and biological activity of 4-alkoxychalcones: potential hydrophobic modulators of P-glycoprotein-mediated multidrug resistance. *Bioorg. Med. Chem*. **7**, 2691-2695

Böttger, V., Böttger, A., Howard, S.F., Picksley, S.M., Chene, P., Garcia-Echeverria, C., Hochkeppel, H.K. & Lane, D.P. (1996) Identification of novel mdm2 binding peptides by phage display. *Oncogene* **13(10)**, 2141-2147

Böttger, A., Böttger, V., Garcia-Echeverria, C., Chene, P., Hochkeppel, H.-K., Sampson, W., Ang, K., Howard, S., Picksley, S. M., and Lane, D. P. (1997) Molecular characterization of the hdm2-p53 interaction. *J. Mol. Biol*. **269**, 744-756

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem*. **72**, 248-254

Branden, C. & Tooze, J. (1999) *Introduction to Protein Structure*. Garland Publ., New York

Brehm, A., Miska, E., Reid, J., Bannister, A. & Kozaurides, T. (1999) The cell cycle-regulating transcription factors E2F-RB. *Br. J. Cancer* **80 (Suppl. 1)**, 38-41

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

Butt, A.J., Firth, S.M., King, M.A. & Baxter, R.C. (2000) Insulin-like growth factor-binding protein-3 modulates expression of Bax and Bcl-2 and potentiates p53-independent radiation-induced apoptosis in human breast cancer cells. *J. Biol. Chem*. **275**, 39174-39181

Cascieri, M.A., Chicchi, G.C., Applebaum, J., Hazes, N.S., Green, B.C. & Bayne, M.L. (1988) Mutants of human insulin-like growth factor I with reduced affinity for the type I insulin-like growth factor receptor. *Biochemistry* **27**, 3229-3233

Chellappan, S.P., Hiebert, S., Mudryj, M., Horowitz, J.M. & Nevins, J.R. (1991) The E2F Transcription Factor Is a Cellular Target For the RB Protein. *Cell* **65**, 1053-1061

Chittenden, T., Livingston, D.M. & Kaelin Jr., W.G. (1991) The T/E1A-binding Domain of the Retinoblastoma Product Can Interact Selectively with a Sequence-specific DNA-binding Protein. *Cell* **65**, 1073-1082

Coligan, J.E., Dunn, B.M., Ploegh, H.L., Speicher, D.W. & Wingfield, P.T. (eds.) (1995) *Current Protocols in Protein Science*. John Wiley& Sons, Inc., USA

Collaborative Computational Project, Number 4. (1994) The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst*. **D50**, 760-763

Cooke, R.M., Harvey, T.S. & Campbell, I.D. (1991) Solution structure of human insulin-like growth factor 1: a nuclear magnetic resonance and restrained molecular dynamics study. *Biochemistry* **30**, 5484-5491

Cress, W.D. & Nevins, J.R. (1996) Use of the E2F Transcription Factor by DNA Tumor Virus Regulatory Proteins. In: Farnham, P.J. (ed.), *Transcriptional Control of Cell Growth: The E2F Gene Family.* Springer Verlag, Berlin, pp.63-78

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

Drenth, J. (1994) *Principles of protein X-ray crystallography*. Springer-Verlag, New York

Dyson, N. (1998) The regulation of E2f by pRB-family proteins. *Genes Dev.* **12**, 2245-2262

El-Deiry, W.S., Kern, S.E., Pietenpol, J.A., Kinzler, K.W. & Vogelstein, B. (1992) Definition of consensus binding site for p53. *Nat. Genet*. **1**, 45-49

Fagan, R., Flint, K.J. & Jones, N. (1994) Phosphorylation of E2F-1 modulates its interactions with the retinoblastoma gene product and the adenoviral E4 19kDa protein. *Cell* **78**, 799-811

Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D. & Kay, L.E. (1994) Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by 15N NMR Relaxation. *Biochemistry* **33**, 5984-6003

Fueyo, J., Gomez-Manzano, C., Yung, W.K.A., Liu, T.J., Alemany, R., McDonnell, T.J., Shi, X., Rao, J.S., Levin, V.A. & Kyritsis, A.P. (1998) Overexpression of E2F-1 in glioma triggers apoptosis and suppresses tumor growth *in vitro* and in *vivo*. *Nature Med*. **4(6)**, 685-690

Giacovazzo, C., Monaco, H.L., Viterbo, D., Scordari, F., Gilli, G., Zanotti, M. & Catti, M. (1992) *Fundamentals of crystallography*. Oxford University Press, Oxford

Gill, R., Wallach, B., Verma, C., Urso, B., Dewolf, E., Grotzinger, J., Murray-Rust, J., Pitts, J., Wollmer, A., Demeyts, P. & Wood, S. (1996) Engineering the C-region of human insuline-like growth factor-1. Implications for receptor binding. *Prot. Eng*. **9,** 1011-1019

Grzesiek, S. & Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc*. **114**, 6291-6293

Hansen, S., Hupp, T.R. & Lane, D.P. (1996) Allosteric regulation of the thermostability and DNA binding activity of human p53 by specific interacting proteins. CRC Cell Transformation Group. *J. Biol. Chem*. **271**, 3917-3924

Haupt, Y., Maya, R., Kazaz, A. & Oren, M. (1997) Mdm2 promotes the rapid degradation of p53. *Nature* **387**, 296-299

Helin, K., Lees, J.A., Vidal, M., Dyson, N., Harlow, E. & Fattaey, A. (1992) A cDNA Encoding a pRB-Binding Protein with Properties of the Transcription Factor E2F. *Cell* **70**, 337-350

Helin, K., Wu, C.-L., Fattaey, A.R., Lees, J.A., Dynlacht, B.D., Ngwu, C. & Harlow, E. (1993a) Heterodimerization of the transcription factors E2F-1 and DP-1 leads to cooperative trans-activation. *Genes Dev*. **7**, 1850-1861

Helin, K., Harlow, E. & Fattaey, A (1993b) Inhibition of E2F-1 transactivation by direct binding of the retinoblastoma protein. *Mol. Cell. Biol*. **13(10)**, 6501-6508

Helin, K. (1998) Regulation of cell proliferation by the E2F transcription factors. *Curr. Op. Genet. Dev*. **8**, 28-35

Hua, Q.X., Shoelson, S.E., Kochoyan, M. & Weiss, M.A. (1991) Receptor binding redefined by a structural switch in a mutant human insulin. *Nature* **354**, 238-241

Huber, H.E., Edwards, G., Goodhart, P.J., Patrick, D.R., Huang, P.S., Ivey-Hoyle, M., Barnett, S.F., Oliff, A. & Heimbrook, D.C. (1993) Transcription factor E2F binds DNA as a heterodimer. *Proc. Natl. Acad. Sci. USA* **90**, 3525-3529

Hwa, V., Oh, Y., Burren, C.P., Choi, W.K., Graham, D.L., Ingermann, A., Kim, H.S., Lopez-Bermejo, A., Minniti, G., Nagalla, S.R., Pai, K., Spagnoli, A., Vorwerk, P., Wanek, D.L.V., Wilson, E.M., Yamanaka, Y., Yang, D.H. & Rosenfeld, R.G. (1999) The IGF binding protein superfamily. In: Rosenfeld, R.G. & Roberts, C.T. (eds), *The IGF system. Molecular Biology, Physiology, and Clinical Applications*. Humana Press, Totowa, pp. 315-327

Imai, Y., Moralez, A., Andag, U., Clarke, J.B., Busby, W.H. & Clemmons, D.R. (2000) Substitutions for hydrophobic amino acids in the N-terminal domains of IGFBP-3 and-5 markedly reduce IGF-I binding and alter their biologic actions. *J. Biol. Chem*. **275,** 18188-18194

Jahnke, W., Baur, M., Gemmecker, G. & Kessler, H. (1995) Improved accuracy of NMR structures by a modified NOESY-HSQC experiment. *J. Magn. Reson*. **106**, 86-88

Johnson, D.G. & Schneider-Broussard, R. (1998) Role of E2F in cell cycle control and cancer. *Front. Biosci*. **3**, 447-458

Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjelgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr*. **A47,** 110-119

Jordan, K.L., Haas, A.R., Logan, T.J. & Hall, D.J. (1994) Detailed analysis of the basic domain of the E2F-1 transcription factor indicates that it is unique among bHLH proteins. *Oncogene* **9**, 1177-1185

Kaelin Jr., W.G., Krek, W., Sellers, W.R., DeCaprio, J.A., Ajchenbaum, F., Fuchs, C.S., Chittenden, T., Li, Y., Farnham, P.J., Blanar, M.A., Livingston, D.M. & Flemington,

E.K. (1992) Expression Cloning of a cDNA Encoding a Retinoblastoma-Binding Protein with E2F-like Properties. *Cell* **70**, 351-364

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

Kaufmann, E. & Knöchel, W. (1996) Five years on the wings of fork head. *Mech. Dev.* **57**, 3-20

Khandwala, H.M., McCutcheon, I.E., Flyvbjerg, A. & Friend, K.E. (2000) The effects of Insulin-Like Growth Factors on Tumorigenesis and Neoplastic Growth. *Endocr. Rev.* **21(3)**, 215-244

Kussie, P.H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levin, A.J. & Pavletich, N.P. (1996) Structure of the MDM2 Oncoprotein Bound to the p53 Tumor Suppressor Transactivation Domain. *Science* **274**, 948-953

Laajoki, L.G., Francis, G.L., Wallace, J.C., Carver, J.A. & Keniry, M.A. (2000) Solution structure and backbone dynamics of long-[Arg3]insulin-like growth factor-I. *J. Biol. Chem*. **275,** 10009-10015

La Fortelle, E. & de Bricogne, G. (1997) Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol*. **276**, 472-494

Lamzin, V.S. & Wilson, K.S. (1993) Automated refinement of protein models. *Acta Cryst*. **D49,** 129-147

Leslie, A.G.W. (1991) Molecular data processing. In Moras, D., Podjarny, A.D. & Thierry,J.C. (eds), *Crystallographic Computing 5*. Oxford University Press, Oxford, UK, pp. 50-61

Liu, T.J., Wang, M., Breau, R.L., Henderson, Y., El-Naggar, A.K., Steck, K.D., Sicard, M.W. & Clayman, G.L. (1999) Apoptosis induction by E2F-1 via adenoviral-mediated gene transfer results in growth suppression of head and neck squamous cell carcinoma cell lines. *Cancer Gene Ther.* **6(2)**, 163-171

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

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

Luo, P. & Baldwin, R.L. (1997) Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **36**, 8413-8421

Marston, F.A.O., Lowe, P.A., Doel, M.T., Schoemaker, J.M., White, S. & Angal, S. (1984) Purification of calf prochymosin (prorennin) synthesized in *Escherichia coli*. *Bio/Technology* **2**, 800-804

Marston, F.A.O. (1986) The purification of eukaryotic polypeptides synthesized in Escherichia coli. *Biochem. J*. **240**, 1-12

Matsushime, H., Roussel, M.F., Ashmun, R.A. & Sherr, C.J. (1991) Colony-stimulating factor 1 regulates novel cyclins during G1 phase of the cell cycle. *Cell* **65**, 701-713

McAlister, M.S.B., Mott, H.R., Van der Merwe, P.A., Campbell, I.D., Davis, S.J. & Driscoll, P.C. (1996) NMR Analysis of Interacting Soluble Forms of the Cell-Cell Recognition Molecules CD2 and CD48. *Biochemistry* **35**, 5982-5991

McPherson, A. (1999) *Crystallization of Biological Macromolecules*. CSHL Press, New York

Müller, H. & Helin, K. (2000) The E2F transcription factor: key regulators of cell proliferation. *Biochim. Biophys. Acta* **1470**, M1-M12

Murray-Rust, J., McLeod, A.N., Blundell, T.L. & Wood, S.P. (1992) Structure and evolution of insulins: implications for receptor binding. *BioEssays* **14,** 325-331

Nevins, J.R. (2001) The Rb/E2F pathway and cancer. *Hum. Mol. Genet.* **10(7)**, 699-703

O'Reilly, D.R., Miller, L. & Luckow, V.A. (1994) *Baculovirus Expression Vectors: A Laboratory Manual.* Oxford University Press, Oxford

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

Qin, X.Q., Livingston, D.M., Kaelin Jr., W.G. & Adams, P.D. (1994) Deregulated transcription factor E2F-1 expression leads to S-phase entry and p53-mediated apoptosis. *Proc. Natl. Acad. Sci. USA* **91(23)**, 10918-10922

Ramirez-Parra, E., Xie, Q., Boniotti, M.B. & Gutierrez, C. (1999) The cloning of plant E2F, a retinoblastoma-binding protein, reveals unique and conserved features with animal G_1/S regulators. *Nucl. Acids Res.* **27(17)**, 3527-3533

Reichel, R., Kovesdi, I. & Nevins, J.R. (1987) Developmental Control of a Promoter-specific Factor That Is Also Regulated by the E1A Gene Product. *Cell* **48**, 501-506

Sambrook, J. & Russell, D.W. (2001) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York

Sato, A., Nishimura, S., Ohkuba, T., Kyogoku, Y., Koyama, S, Kobayashi, M., Ysuda, T. & Kobayashi, Y. (1993) Three-dimensional structure of human insulin-like growth factor-1 (IGF-1) determined by ^{1}H NMR and distance geometry. *Int. J. Pept. Protein Res*. **41**, 433-440

Schagger, H. & von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem*. **166(2)**, 368-379

Sekine, M., Ito, M., Uemukai, K., Maeda, Y., Nakagami, H. & Shinmyo, A. (1999) Isolation and characterization of the E2F-like gene in plants. *FEBS Lett.* **460(1)**, 117-122

Shan, B., Zhu, X., Chen, P.-L., Durfee, T., Yang, Y., Sharp, D. & Lee, W.-H. (1992) Molecular Cloning of Cellular Genes Encoding Retinoblastoma-Associated Proteins: Identification of a Gene with Properties of the Transcription Factor E2F. *Mol. Cell. Biol.* **12(12)**, 5620-5631

Sheldrick, G. (1991) Tutorial on automated Patterson interpretation to find heavy atoms. In Moras, D., Podjarny, A.D. & Thierry, J.C. (eds), *Crystallographic Computing 5*. Oxford University Press, Oxford, UK, pp. 145-157

Sherr, C.J. (1996) Cancer Cell Cycles. *Science* **274**, 1672-1677

Sherr, C.J. (2000) The Pezoller Lecture: Cancer Cell Cycles Revisited. *Cancer Res.* **60**, 3689-3695

Shuker, S.B., Hajduk, P.J., Meadows, R.P. & Fesik, S.W. (1996) Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **274**, 1531-1534

Sladek, T.L. (1997) E2F transcription factor action, regulation and possible role in human cancer. *Cell Prolif.* **30**, 97-105

Slansky, J.E. & Farnham, P.J. (1996) Introduction to the E2F Family: Protein Structure and Gene Regulation. In: Farnham, P.J. (ed.), *Transcriptional Control of Cell Growth: The E2F Gene Family.* Springer Verlag, Berlin, pp. 1-31

Stoll, R., Renner, C., Hansen, S., Palme, S., Klein, C., Belling, A., Zeslawski, W., Kamionka, M., Rehm, T., Mühlhahn, P., Schumacher, R., Hesse, F., Kaluza, B., Voelter, W., Engh, R.A. & Holak, T.A. (2001) Chalcone Derivatives Antagonize Interactions between the Human Oncoprotein MDM2 and p53. *Biochemistry* **40**, 336-344

Strom, D.K., Cleveland, J.L., Chellappan, S., Nip, J. & Hiebert, S.W. (1998) E2F-1 and E2F-3 are functionally distinct in their ability to promote myeloid cell cycle progression and block granulocyte differentiation. *Cell Growth Differ.* **9(1)**, 59-69

Tao, Y., Kassatly, R.F., Cress, W.D. & Horowitz, J.M. (1997) Subunit composition determines E2F DNA-binding site specificity. *Mol. Cell. Biol*. **17(12)**, 6994-7007

Terasawa, H., Kohda, D., Hatanaka, H., Nagata, K., Higashihashi, N., Fujiwara, H., Sakano,K. & Inagaki, F. (1994) Solution structure of human insulin-like growth factor II; recognition sites for receptors and binding proteins. *EMBO J.* **13***,* 5590-5597

Torres, A.M., Forbes, B.E., Aplin, S.E., Wallace, J.C., Francis, G.L. & Norton, R.S. (1995) Solution structure of human insulin-like growth factor II. Relationship to receptor and binding protein interactions. *J. Mol. Biol*. **248**, 385-401

Voet, D. & Voet, J.G. (1995) *Biochemistry.* John Wiley & Sons, New York

Wetterau, L.A., Moore, M.G., Lee, K.-W., Shim, M.L. & Cohen, P. (1999) Novel Aspects of the Insulin-like Growth Factor Binding Proteins. *Mol. Genet. Metabol.* **68**, 161-181

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore

Yee, A.S., Raychaudhuri, P., Jakoi, L. & Nevins, J.R. (1991) The Adenovirus-Inducible Factor E2F Stimulates Transcription after Specific DNA Binding. *Mol. Cell. Biol.* **9(2)**, 578-585

Zheng, N., Fraenkel, E., Pabo, C.O. & Pavletich, N.P. (1999) Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F-DP. *Genes Dev*. **13**, 666-674