Technische Universität München

Max-Planck-Institut für Biochemie

Abteilung Strukturforschung

Biologische NMR-Arbeitsgruppe

Structural and biochemical characterization of cell cycle regulatory proteins and associated signaling receptors

Kinga Dorota Rehm

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

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ. - Prof. Dr. Chr. F. W. Becker

Prüfer der Dissertation: 1. Univ. - Prof. Dr. M. Sattler

2. Univ. - Prof. Dr. M. Groll

Die Dissertation wurde am 28.09.2009 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 26.11.2009 angenommen.

To my Parents...

Acknowledgements

I would like to express my gratitude to all those who gave me the possibility to complete this thesis:

To Professor Michael Sattler, for being my Doktorvater.

To my supervisor Tad A. Holak, for giving me the chance to work in his group and for his support.

To Professor Olaf Stemmann, for giving me the opportunity to participate in his project, a lot of practical advice, scientific support and access to his unpublished materials.

To Professor Adam Dubin.

To Ola.

To all the colleagues, who helped me solving daily lab problems, shared their experience and created a friendly atmosphere in the lab.

Last but not least to Till.

Publications

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

Ulli Rothweiler, Anna Czarna, Lutz Weber, Grzegorz M. Popowicz, **Kinga Brongel**, Kaja Kowalska, Michael Orth, Olaf Stemmann and Tad A. Holak
NMR screening for lead compounds using tryptophan-mutated proteins
J. Med. Chem. 2008, *51*: 5035–5042

Bernd Mayer, Michael Orth, **Kinga Rehm**, Kay Hofmann, Tad A. Holak, and Olaf Stemmann.
Shugoshin is a Mad1/Cdc20-like interactor of Mad2
Manuscript under preparation to be submited to EMBO J.

Contents

1 Introduction

The various types of malignant cancers known today are still one of the leading causes of death in the industrialized world, second only to cardiovascular diseases. Many of the aspects of cancer development, the cell cycle and the regulation of tissue growth remain unknown and are subject to intense research. The function of the multitude of proteins involved in the cell cycle is closely related to their three dimensional structure. Knowledge of a protein's structure and changes in the structure, that is its dynamics, is the key to the understanding of its function – and also its malfunction. Thus protein structure research is one of the central issues in life sciences today. This work is a small contribution to the few answers in an ocean of open questions.

Scope of this work

This thesis is a collection of three projects conducted at the Department of Structural Research of the Max Planck Institute of Biochemistry in recent years. In each of theses projects the structural and biochemical characterization of a protein important in cancer development was attempted. Details on the biological background of each project are given in the introductions to each chapter. The used materials and methods are explained in chapter 2.

In chapter 3 work on the IGF-1 receptor, the specific cell surface receptor for IGF, involved in normal growth and development is described. A precise model of the IGF-1R activation by IGF might be created based on x-ray structures of the interacting proteins and information on their dynamic behaviour from NMR studies. The trials of expression and purification of IGF-1R fragments for further structural research are presented.

A crucial step in the cell cycle is the nuclear division. Currently research focuses on the role of Shugoshin and Mad2 and their potential interaction for the regulation of sister chromatids separation. The elucidation of the three dimensional structure of the complex would be a crucial step in understanding

the mechanisms of the mitotic checkpoint. Attempts to crystallize Mad2 in complex with Sgo fragments are discussed in chapter 4 as well as NMR and ITC binding studies between these two proteins.

Frequent deregulation of CDK2 in cancers has lead to the search of different compounds that can specifically inhibit this kinase and until today a large number of different inhibitors of CDK2 are known. In chapter 5 the attempts to determine the molecular basis of action of a new small molecule inhibitor for CDK2 by means of x-ray crystallography are reported. CDK2 – inhibitor interaction was also chosen to illustrate a novel NMR-based drug screening method which permits rapid screening by monitoring a simple 1D proton NMR signal of the NH side chain ($^{N}H^{\varepsilon}$) of the tryptophans.

2 Materials and laboratory techniques

2.1 Materials

2.1.1 Chemicals

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

2.1.2 Enzymes and antibodies

BSA	New England BioLabs (USA)
CIP	New England BioLabs (USA)
BamH I	New England BioLabs (USA)
EcoR I	New England BioLabs (USA)
Hind III	New England BioLabs (USA)
Nde I	New England BioLabs (USA)
Xho I	New England BioLabs (USA)
Not I	New England BioLabs (USA)
Nco I	New England BioLabs (USA)
Xba I	New England BioLabs (USA)
Pfu turbo DNA Polymerase	Stratagene (USA)
Pfu DNA Polymerase	Fermentas (Lithuania)
Phusion DNA Polymerase	Finnzymes (Finland)
T4 DNA Ligase	New England BioLabs (USA)
Dpn I	Stratagene (USA)
Xa Factor	Novagen (Canada)
Enterokinase	Novagen (Canada)
Senp Protease	Core Facility MPI (Germany)
TAGzyme	Qiagen (Germany)
TEV Protease	Sigma (USA)
Anti His antibodies (mouse)	Santa Cruz biotech (USA)
Goat anti mouse antibodies	Santa Cruz biotech (USA)

2.1.3 Kits and reagents

QIAquick PCR Purification Kit	Qiagen (Germany)
QIAquick Gel Extraction Kit	Qiagen (Germany)
QIAprep Spin Miniprep Kit	Qiagen (Germany)
QuikChange Site-Directed Mutagenesis Kit	Stratagene (USA)
Pre-Crystallization Test (PCT)	Hampton Research (USA)
Complete Protease Inhibitor Cocktail	Roche (Germany)
pET LIC cloning Kits	Novagen (Canada)

2.1.4 Oligonucleotides

Oligonucleotides used for cloning were provided by Metabion, or MWG (Germany).

2.1.5 *E. coli* and *P. pastoris* strains

- Cloning strains

XL1-Blue	Stratagene (USA)
One Shot TOP10	Invitrogen (Holland)
DH5α	Novagen (Canada)

- Protein expression *E. coli* strains

BL21 Star(DE3)	Invitrogen (Holland)
BL21 Star(DE3) pLysS	Invitrogen (Holland)
Rosetta(DE3)	Novagen (Canada)
RosettaGami(DE3)	Novagen (Canada)
BL21(DE3) CodonPlus	Novagen (Canada)
BL21(DE3) - pGKJE8	Novagen (Canada)
Tuner	Novagen (Canada)
ArcticExpress RIL- pRARE	Novagen (Canada)

- Protein expression *P. pastoris* strains

KM71H	Invitrogen (Holland)
SDM1168H	Invitrogen (Holland)

2.1.6 Vectors

pET 44 Ek/LIC	Novagen (Canada)
pET 46 Ek/LIC	Novagen (Canada)
pETM-20	Novagen (Canada)
pETM-50	Novagen (Canada)
pETM-41	Novagen (Canada)
pET20b	Novagen (Canada)
pET40b	Novagen (Canada)
pET 28a	Novagen (Canada)
pQE 80	Qiagen (Germany)
pRFSDuet-1	Novagen (Canada)
pMal-Tev2-FA	Core Facility MPI (Germany)
pET28M-Sumo1-GFP	Core Facility MPI (Germany)
pProExHTa	Invitrogen (Holland)
pICZB	Invitrogen (Holland)
pICZαC	Invitrogen (Holland)

2.1.7 Cell growth media and stocks

2.1.7.1 Media

- **Luria Bertani (LB)** medium:

For 1 liter LB: 10 g bacto tryptone
5 g bacto yeast extract
10 g sodium chloride

pH was adjusted to 7.0. For the preparation of agar plates the medium was supplemented with 15 g agar. The mixture was autoclaved.

- **Minimal medium (MM)** for uniform enrichment with ^{15}N:

For 1 liter MM: 0.5 g NaCl
1.3 ml trace elements solution
1 g citric acid monohydrate
36 mg ferrous citrate

4.02 g KH_2PO_4
7.82 g K_2HPO_4 x $3H_2O$
1 ml Zn-EDTA solution
1 g NH_4Cl or $^{15}NH_4Cl$

pH was adjusted to 7.0 with NaOH, the mixture was autoclaved, upon cooling separately sterilized solutions were added: 15 ml 40% glucose, 560 µl thiamin, antibiotics, 2 ml $MgSO_4$ stock.

- **Yeast extract Peptone (YP)** medium:

For 1 liter YP: 20 g bacto tryptone
10 g bacto yeast extract
50 ml 40% glucose

pH was adjusted to 6.5. For the preparation of agar plates the medium was supplemented with 20 g agar. The mixture was autoclaved.

- **Buffered Minimal Methanol Histidine (BMMH) medium**:

For 1 liter BMMH: 700 ml H_2O
100 ml 5% methanol
100 ml phosphate buffer
100 ml YNB
10 ml Hisidine stock solution

Every solution was separately sterilized and then mixed.

2.1.7.2 Stock solutions

- **Stock solution of glucose**

40 g of glucose were dissolved in distilled water to the final volume of 100 ml and autoclaved.

- **Stock solution of ampicillin**

100 mg/ml of ampicillin were dissolved in distilled water. The stock solution was sterile filtrated and stored in aliquots at -20°C until used.

- **Stock solution of kanamycin**

50 mg/ml of kanamycin were dissolved in distilled water. The stock solution was sterile filtrated and stored in aliquots at -20°C until used.

- **Stock solution of chloramfenicol**

50 mg/ml of chloramfenicol were dissolved in ethanol. The stock solution was sterile filtrated and stored in aliquots at -20°C until used.

- **Stock solution of streptomycin**

50 mg/ml of streptomycin were dissolved in distilled water. The stock solution was sterile filtrated and stored in aliquots at -20°C until used.

- **Stock solution of tetracycline**

12.5 mg/ml tetracycline were dissolved in 50% ethanol. The stock solution was sterile filtrated and stored in aliquots at -20°C until used.

- **Stock solution of IPTG**

A sterile 1 M stock of IPTG in distilled water was prepared and stored in aliquots at -20°C

- **Stock solution of thiamin**

1% Thiamin in deionised H_2O was sterile filtrated and stored at 4°C until used.

- **Stock solution of $MgSO_4$**

A sterile 1 M stock of $MgSO_4$ in distilled water was prepared and stored at 4°C until used

- **Zn-EDTA solution**

5 mg/ml EDTA

8.4 mg/ml $Zn(Ac)_2$

- **Trace elements solution**

2.5 g/l H_3BO_3

2.0 g/l $CoCl_2 \times H_2O$

1.13 g/l $CuCl_2 \times H_2O$

9.8 g/l $MnCl_2 \times 2H_2O$

2.0 g/l $Na_2MoO_4 \times 2H_2O$

pH was lowered with citric acid or HCl.

- **Phosphate buffer**

For 1 liter: 132 ml 1M K_2HPO_4

868 ml 1M KH_2PO_4

- **Histidine stock solution**

400 mg Histidine in 100 ml distilled water filtrated and stored at 4℃ until used.

- **Yeast Nitrogen Base (YNB)**

6.7 g in distilled water, filtrated.

- **1M sorbitol solution**

A sterile 1 M stock of IPTG in distilled water was prepared.

2.1.8 Solutions for chemically competent *E. coli* cells

Buffer A	100 mM $MgCl_2$ x 6H2O
Buffer B	100 mM $CaCl_2$ –15% glycerol

2.1.9 Antibiotics

Ampicillin, sodium salt	Sigma (USA)
Kanamycin, monosulfate	Sigma (USA)
Chloramphenicol	Sigma (USA)
Streptomycin	Sigma (USA)
Tetracycline	Sigma (USA)
Zeocin	Invitrogen (Holland)

2.1.10 Isotopically enriched chemicals

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

2.1.11 Buffer for DNA agarose gel electrophoresis

50X TAE buffer	for 1 l:
40 mM Tris-acetate	242 g of Tris base
1 mM EDTA	100 ml of 0.5 M EDTA (pH 8.0)
Glacial acetic acid	57.1 ml

2.1.12 Protein purification – buffers

- **PBS** 140 mM NaCl
  2.7 mM KCl
  10 mM Na_2HPO_4
  1.8 mM KH_2PO_4
  0.05% NaN_3
  pH 7.3

- **Standard buffers for immobilized metal-chelate chromatography (IMAC) under native conditions**

Binding buffer 50 mM NaH_2PO_4
300 mM NaCl
10 mM imidazole
pH 8.0

Wash buffer 50 mM NaH_2PO_4
300 mM NaCl
20 mM imidazole
pH 8.0

Elution buffer 50 mM NaH_2PO_4
300 mM NaCl
250 mM imidazole
pH 8.0

- **Standard buffers for IMAC under denaturing conditions**

Buffer A (binding buffer) 6 M guanidinium chloride
100 mM NaH_2PO_4 x H_2O
10 mM Tris
10 mM β-mercaptoethanol
pH 8.0

Buffer B (wash buffer) 6 M guanidinium chloride
100 mM NaH_2PO_4 x H_2O
10 mM Tris
10 mM β-mercaptoethanol
pH 6.5

Buffer C (elution buffer) 6 M guanidinium chloride
pH 3.0

- **Buffer for His6-xMad2-xSgo copurification**

Lysis buffer 50 mM NaH_2PO_4
500 mM NaCl
20 mM imidazole
5 mM DTT
5 mM EDTA
pH 7.4

- **Buffers for His6-xMad2 and xMad2-xSgo complex purification**

Lysis buffer 50 mM NaH_2PO_4
400 mM NaCl
10 mM imidazole
5 mM DTT
pH 7.4

Wash buffer 50 mM Hepes/KOH
100 mM NaCl
20 mM imidazole
0.1% Tween-20
5 mM β-mercaptoethanol
pH 7.0

Elution buffer 50 mM Hepes/KOH
100 mM NaCl
200 mM imidazole
0.1% Tween-20
5 mM β-mercaptoethanol
pH 7.0

- **Other buffers used for immobilized metal-chelate chromatography (IMAC) under native conditions of different IGF-1R constructs**

Binding buffer 50 mM NaH_2PO_4
200 arginine
10 mM imidazole
pH 8.0

Wash buffer 50 mM NaH_2PO_4
200 mM arginine
30 mM imidazole
pH 8.0

Elution buffer 50 mM NaH_2PO_4
150 mM NaCl
250 mM imidazole
pH 7.5

- **Refolding buffers**

Refolding buffer 200 mM arginine HCl
1 mM EDTA
100 mM Tris
2 mM red GSH
2 mM ox GSH
10% (v/v) glycerol
pH 8.4

Refolding buffer	200 mM arginine HCl 1 mM EDTA 100 mM Tris 5 mM red GSH 1 mM ox GSH 10% (v/v) glycerol pH 8.4
Refolding buffer	200 mM arginine HCl 1 mM EDTA 100 mM Tris 10 mM cysteamine 5 mM cystamine 10% (v/v) glycerol pH 8.4
Refolding buffer	500 mM arginine HCl 1 mM EDTA 100 mM Tris 5 mM red GSH 1 mM ox GSH 10% (v/v) glycerol pH 8.4

- **Protease buffers**

Buffer X (Factor Xa cleavage buffer)	50 mM Tris 100 mM NaCl 4 mM $CaCl_2$ 0.05% NaN_3 pH 8.0

Buffer TEV (TEV cleavage buffer) 50 mM NaH_2PO_4
150 mM NaCl
0.05% NaN_3
pH 7.5

Buffer EK (Enterokinase cleavage buffer) 20 mM Tris
100 mM NaCl
2 mM CaCl2
0.01% NaN3
pH 7.5

2.1.13 Reagents and buffers for the SDS-PAGE

Anode buffer (+): 200 mM Tris pH 8.9

Cathode buffer (-): 100 mM Tris pH 8.25
100 mM tricine
0.1% SDS

Separation buffer: 1 M Tris pH 8.8
0.3% SDS

Stacking buffer: 1 M Tris pH 6.8
0.3% SDS

Separation acrylamide: 48% acrylamide
1.5% bis-acrylamide

Stacking acrylamide: 30% acrylamide
0.8% bis-acrylamide

Pouring polyacrylamide gels:

Separation gel: 1.675 ml H_2O
2.5 ml separation buffer
2.5 ml separation acrylamide
0.8 ml glycerol
25 µl APS
2.5 µl TEMED

Intermediate gel: 1.725 ml H_2O
1.25 ml separation buffer
0.75 ml separation acrylamide
12.5 µl APS
1.25 µl TEMED

Stacking gel: 2.575 ml H_2O
0.475 ml stacking buffer
0.625 ml stacking acrylamide
12.5 µl 0.5 M EDTA, pH 8.0
37.5 µl APS
1.9 µl TEMED

2.1.14 Protein visualization

Coomassie-blue solution: 45% ethanol
10% acetic acid

Destaining solution: 5% ethanol
10% acetic acid

2.1.15 Reagents and buffers for western blots

Transfer buffer	10 mM CAPS 20% Methanol 25 mM Tris 192 mM glycine pH 11.0
Alkaline phosphatase buffer	100 mM Tris 100 mM NaCl 5 mM $MgCl_2$ pH 9.5
Wash buffer	10 mM Tris 150 mM NaCl 0.05% Tween20 pH 8.0
1^{st} antibody solution	1:2000 diluted in the wash buffer
2^{nd} antibody solution (linked to alkaline phosphatase)	1:2000 diluted in alkaline phosphatase buffer
Substrate for alkaline phosphatase	BCIP (Sigma); dissolve 1 tablet in 10 ml of water

2.1.16 Protein and nucleic acids markers

Prestained Protein Marker	New England BioLabs (USA)
100 BP DNA marker	New England BioLabs (USA)
1Kb DNA marker	New England BioLabs (USA)
Broad Range (6-175 kDa) 1 kb DNA-Leiter	Peqlab (Germany)

2.1.17 Peptides

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

2.1.18 Apparatus

2.1.18.1 ÄKTA explorer 10 purification system

ÄKTAexplorer 10 (Amersham Biosciences (Sweden)) is a complete fully automated liquid chromatography system consisting of a separation unit, fraction collector Frac-901 and a personal computer running UNICORN software v.3.1 to control the unit. The system enables to monitor light absorption at 3 wavelengths simultaneously, in the range 190-700 nm; pH; conductivity; temperature; flow rate and pressure values. High performance pumps make it possible to apply flow rates in the range of 0.01-10 ml/min and gradients. Samples were applied onto a 10 or 50 ml superloop (Amersham Biosciences, Sweden), which was filled manually. UV absorption was monitored at wavelengths 255, 270 and 280nm. As the pH of the buffers was known and constant, pH was not monitored. Runs were performed in a cold cabinet, at 4°C.

2.1.18.2 Chromatography equipment, columns and media

ÄKTA explorer 10	Amersham Pharmacia (Sweden)
Peristaltic pump P-1	Amersham Pharmacia (Sweden)
Fraction collector RediFrac	Amersham Pharmacia (Sweden)
Recorder REC-1	Amersham Pharmacia (Sweden)
UV flow through detector UV-1	Amersham Pharmacia (Sweden)
BioloLogic LP System	Biorad (USA)
HiLoad 26/60 Superdex S75pg	Amersham Pharmacia (Sweden)
HiLoad 16/60 Superdex S75pg	Amersham Pharmacia (Sweden)
HiLoad 16/60 Superdex S200pg	Amersham Pharmacia (Sweden)
HiLoad 10/30 Superdex S75pg	Amersham Pharmacia (Sweden)
HiLoad 10/30 Superdex S200pg	Amersham Pharmacia (Sweden)
NiNTA-agarose	Qiagen (Germany)

2.1.18.3 NMR spectrometer

600 MHz Bruker (Germany)

2.1.18.4 Other apparatus

Autoclave	Bachofer Reutlingen (Germany)
Balances PE 1600 AE 163	Mettler (Germany)
BIO-RAD MINI-PROTEAN II gel apparatus	Bio-Rad Laboratories (UK)
Centrifuge 3K15	Sigma (Germany)
Centrifuge 5414	Eppendorf (Germany)
Centrifuge Avanti J-30I	Beckman (USA)
Centrifuge Microfuge R	Beckman (USA)
Centrifuges: J-6M/E and Avanti J-30I	Beckman (UK)
Chambers for SDS PAGE and Western blotting	BioRad (Germany)
Ice machine Scotsman AF 30	Frimont Bettolino (Italy)
Incubator shaker	Lab-Therm Kühner (Switzerland)
Microcalorimeter VP/ITC	Microcal USA
Magnetic stirrer Heidolph M2000	Bachofer Reutlingen (Germany)
MARresearch image plates, mar345	MARresearch (Germany)
PCR machine: Mastercycler personal	Eppendorf (Germany)
pH-meter pHM83 Radiometer	Copenhagen (Denmark)
Pipettes 2.5µl, 10µl, 20µl, 200µl, 1000µl	Eppendorf (Germany)
Quarz cuvettes	QS Hellma (Germany)
Shaker	Adolf-Kühner AG (Switzerland)
Sonicator: Sonifer 250	Branson (USA)
Spectrophotometer	Amersham Pharmacia (Germany)
Spectrophotometer Ultraspec 3100 pro	AmershamBiosciences (Sweden)
Western blot apparatus (semidry)	MPI for Biochemistry

2.2 Laboratory methods and principles

The most common protocols for procedures or formulations of solutions are taken directly or adapted from 'Molecular Cloning' (Sambrook 2001). In modifying or designing protocols manuals delivered with the enzymes and chemicals were also very helpful and provided necessary information.

2.2.1 Construct design and choice of the expressions system

The proper design of the protein construct is essential for the protein expression and successful crystallization. Unstructured and flexible fragments of proteins usually inhibit crystallization or turn out crystals of low quality. Optimized protein constructs result in well-defined, folded and stable proteins, and at the same time guarantee its biological activity. Designing of the constructs may be based on previously published literature and secondary structure prediction with the help of bioinformatics tools. Determination of the stable and folded constructs requires employing various techniques like for example: limited proteolysis, protein sequencing, mass spectrometry, and NMR spectroscopy (Rehm et al., 2002).

The key to successful expression of the recombinant proteins is the choice of the proper expression system (Makrides, 1996). Many systems are available and each offers particular advantages and disadvantages that may be pertinent to the specific goal. In the majority of cases, the popular well-described systems are suitable choices: expression in *E. coli,* the baculovirus system in insect cells, yeast systems (*Saccharomyces cerevisiae* and *Pichia pastoris*) and mammalian CHO cells. The desired quality and quantity of the protein is determining the choice of expression. With respect to quality, the most important issue is the inclusion or exclusion of post-translational modifications of the protein (e.g. glycosylation). The current major choices of systems for expression of glycosylated proteins are yeasts, insect cells, and mammalian cells. The second important criterion is the amount of protein required. Expression in *E. coli* is a simple and economic way to obtain relatively high amounts of protein of interest.

Unfortunately expression in *E. coli* often results in accumulation of the foreign protein as inclusion bodies. Poor overall protein recovery resulting from poor folding yield can be compensated by increasing the scale of expression or choosing the appropriate purification strategy. When protein bearing post-translational modifications are needed, yeast expression (e.g. *Pichia pastoris*) can be successfully used. In addition yeast allows expression and secretion of multidomain proteins (>50 kDa) that are often poorly expressed in *E. coli* (Gray and Subramanian, 2001).

In this work mostly vectors from the pET series were used for the expression of proteins in *E. coli* (pET system manual, Novagen, 2003; Amersham Pharmacia, 2003). The presence of a 6-histidine fusion peptide (His-Tag) makes use of immobilized metal chromatography (IMAC) enabling rapid purification. Large fusion tags, such as SUMO, thioredoxin or maltose binding protein can help to increase solubility and promote proper folding of recombinant fusion partners. Later the removal of the fusion tag can be achieved with the aid of a specific restriction protease. Vectors with specific protease sites are commercially available otherwise they could be inserted using PCR.

2.2.2 DNA techniques

2.2.2.1 PCR

A polymerase chain reaction was employed to amplify desired DNA fragments and genes, introduce restriction sites, STOP codons and sequences encoding restriction protease cleavage sites. The primers were prepared according to standardized principles regarding the length, GC-content, melting temperature and occurrence of secondary structures of the hairpin type. All primers used for cloning and mutagenesis are listed in Table 2.1. Two different kinds of recombinant thermostable DNA polymerases were used, each operating at slightly different conditions:

	Melting temp.	Annealing temp.	Synthesis temp.
Phusion HF	98°C	primer melting temperature -5°C	72°C
Pfu Turbo	95°C	primer melting temperature -5°C	72°C

The reaction was prepared as follows:

10 x buffer	5 µl
dNTP 100mM	0.8 µl
template DNA	50 or 100 ng
sens_primer 100 ng/µl	2.5 µl
antisens_primer 100 ng/µl	2.5 µl
nuclease free water	up to 49 µl total volume
Pfu Turbo polymerase	1 µl

For the *ligation independent cloning* the PCR amplification of the gene of interest must provide the compatible overhang at both ends (LIC cloning kit). Following are the sequences of the overhangs for pET41 and 46 LIC/Ek:

Forward primer 5' – GACGACGACAAGAT – 3'

Reverse primer 5' – GAGGAGAAGCCCGGT – 3'

Table 2.1. List of primers used in this work

No	Name	Nucleotide sequence
1	IGF1R31**F**	gac gac gac aag ata gaa atc tgc ggg cca ggc atc gac atc cgc
2	IGF1R213**R**	gag gag aag ccc ggt *tta* ttt ctg gca gcg gtt tgt ggt cca gca gcg
3	IGF1R306**R**	gag gag aag ccc ggt *tta* ctc ctg cat gca ctc gcc gtc gtg gat cac
4	IGF1R331**R**	gag gag aag ccc ggt *tta* gac ctt cgg gca agg acc ttc aca agg g
5	IGF1R490**R**	gag gag aag ccc ggt *tta* act ttc aca gga ggc tca cac ccc gtt gtt
6a	IGF1R306**R**CT1	cag gaa att ctc aaa gac ttt ctc ctg cat gca ctc gc
6b	IGF1R306**R**CT2	ggg cac gaa gat gga gtt gtg cag gaa att ctc aaa gac ttt
6c	IGF1R306**R**CT3	ttc agg tct ggg cac gaa gat gga gtt gtg
6d	IGF1R306**R**CT4	gag gag aag ccc ggt *tta* ttc agg tct ggg cac gaa
7	IGF1R31**F**NcoI	gga cca tgg aaa tct gcg ggc cag g
8	IGF1R306**R**BamHI	cgt gga tcc *tta* ctc ctg cat gca ctc gc
9	IGF1R331**R**BamHI	cca gga tcc *tta* gac ctt cgg gca agg acc
10	IGF1R490**R**BamHI	cga gga tcc *tta* act ttc aca gga ggc
11	IGF1RαC**F**EcoRI	ggc gaa ttc a gaa atc tgc ggg cca gg

12	IGF1RB**F**EcoRI	ggc gaa ttc gaa atg gaa atc tgc ggg cca gg
13	IGF1RαC**R**CTNotI	gc ggc cgc acg acc ttc gat ttc agg tct ggg cac gaa
14	IGF1RαC**R**331NotI	gc ggc cgc acg acc ttc gat gac ctt cgg gca agg acc
15	IGF1RαC**R**490NotI	gc ggc cgc acg acc ttc gat act ttc aca gga ggc
16	IGF1RB**R**CTXbaI	cgc tct ag acg acc ttc gat ttc agg tct ggg cac gaa
17	IGF1RB**R**331XbaI	cgc tct ag acg acc ttc gat gac ctt cgg gca agg acc
18	IGF1RB**R**490XbaI	cgc tct ag acg acc ttc gat act ttc aca gga ggc
19	xSgoS105**F**OS	cag cag gat ccc agc aag gaa gag ata ctg
20	xSgoS51**F**OS	cgc cca gga tcc aca gcg atc aag aac tat aaa
21	xSgoN188**R**OS	gta gcg gcc gct agt ttg gta tca cat ttt tgg gt
22	xMad2**F**NdeOS	ccc agc ata tgg cgg ggc agc taa ca
23	xMad2**R**XhoOS	ggc cgg ctc gag tta gaa agt gct tat ctt ttt g
24	CDK2S53W**F**	gcccagtactgcc atc cga gag atc tgg ctg ctt aag gag ctt aac cat cc
25	CDK2S53W**R**	ggatggttaagctc ctt aag cag cca gat ctc tcg gat ggc agt act ggg c
26	CDK2H84W**F**	c ctg gtt ttt gaa ttt ctg tgg caa gat ctc aag aaa ttc atg gat gcc
27	CDK2H84W**R**	ggc atc cat gaa ttt ctt gag atc ttg cca cag aaa ttc aaa aac cag g
28	CDK2A93W**F**	c caa gat ctc aag aaa ttc atg gat tgg tct gct ctc act ggc att cc
29	CDK2A93W**R**	gg aat gcc agt gag agc aga cca atc cat gaa ttt ctt gag atc ttg g

***F** – forward; **R** – reverse; stop codons in italics; mutations are coloured red.*

2.2.2.2 Digestion with restriction enzymes

1-2 units of each restriction enzyme were used per 1 µg of plasmid DNA to be digested. The digestion was performed in a buffer specified by the manufacturer at the optimal temperature (37℃) for 1-3 h. The fragment's ends that occurred after digestion were cohesive. To eliminate the possibility of plasmid recirculation (possible when double-digestion does not occur with 100% efficiency), 5'-ends of a vector were dephosphorylated using calf intestine phosphatase (CIP). CIP treatment was performed with 1 unit of enzyme per 3 µg of plasmid DNA, at 37 ℃ for 1 h.

2.2.2.3 T4 DNA polymerase digestion for ligation independent cloning

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

0.2 pmol purified PCR product in up to 14.6 µl water

2 µl 10X T4 DNA polymerase buffer

2 µl dATP 25 mM

1 µl DTT 100 mM

0.4 µl T4 DNA Polymerase

2.5 U/µl (0.5 unit per 0.1 pmol PCR product) (LIC – qualified)

up to 20 µl nuclease free water

The final concentration of insert is 0.01 pmol/µl. The reaction was incubated at 22°C for 30 min and subsequently inactivated by heating up to 75°C for 20 min.

2.2.2.4 Purification of PCR and restriction digestion products

DNA obtained from restriction digestion, phosphatase treatment or PCR was purified from primers, nucleotides, enzymes, buffering substances, mineral oil, salts, agarose, ethidium bromide, and other impurities, using a silica-gel column (QIAquick PCR Purification Kit, Qiagen). The QIAquick system uses a simple bind-wash-elute procedure. A binding buffer was added directly to the PCR sample or other enzymatic reaction, and the mixture was applied to the spin column. Nucleic acids adsorbed to the silica-gel membrane in the high-salt conditions provided by the buffer. Impurities and short fragments of single or double-stranded DNAs were washed away and pure DNA was eluted with a small volume of 10 mM Tris pH 8.0 or water.

2.2.2.5 Ligation

The ligation of digested and purified inserts and vectors was performed according to the protocol described in the T4 DNA Ligase instruction (New England BioLabs).

The ligation mixture contained (20 µl):

insert	16 µl (0.15 µM)
10 x T4 DNA ligase reaction buffer	2 µl
vector	1 µl
T4 DNA Ligase	1 µl

2.2.2.6 Ligation independent cloning

The ligation reaction was assembled with following components:

1 µl the LIC vector (approximately 0.01 pmol of LIC vectors)

2 µl (0.02 pmol) T4 DNA polymerase-treated LIC insert

The reaction was incubated at 22°C for 5 min and then, after addition of 1 µl of 25 mM EDTA, again for 5 min at 22°C. The resulting 4 µl of the ligation reaction mixture was used for the transformation of *E. coli* competent cells.

2.2.2.7 Screening of positive colonies

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

2.2.2.8 Isolation of plasmid DNA

The preparation of plasmid DNA from *E. coli* was carried out using dedicated plasmid purification kits from Qiagen. The kits employ a standard alkalic lysis of the precipitated bacteria in the presence of RNAse and a strong ionic detergent, SDS, followed by neutralization/DNA renaturation with acetate. For purification, a crude cell lysate is loaded onto a silica gel column, washed with an ethanol-containing buffer, and eluted in a small volume, yielding up to 20 µg of the plasmid DNA.

2.2.2.9 DNA sequencing

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

2.2.2.10 Mutagenesis

Site directed mutagenesis of CDK2 was performed with PCR, using enzymes and instructions supplied in the QuikChange Site-Directed Mutagenesis Kit (Stratagene). The mutagenic oligonucleotide primers were designed according to suggestions provided by the manufacturer. The desired mutation was in the middle of the primer with ~ 20-30 bases of a wild type sequence on both sides (Table 2.1). Vector pET 28a containing the copy of a gene encoding CDK2 wild type was used as DNA template. High concentration of the template DNA and low number of PCR cycles, combined with high accuracy and fidelity of highly processive DNA polymerase *Pfu* Turbo, minimizes the occurrence of unwanted mutations.

The mutagenic PCR reaction mixture contained:

10 × reaction buffer	5 µl
dNTP mix	1 µl
plasmid (5 ng/µl)	1 µl (10 ng)
oligonucleotide primer F	2 µl (125 ng)
oligonucleotide control primerF	2 µl (125 ng)
PfuTurbo DNA polymerase (2.5 U/µl)	1 µl
milli-Q to a final volume of 50 µl	38 µl

PCR cycling parameters:

denaturation:	95°C, 1’
denaturation:	95°C, 30’’
annealing:	55°C, 1’
synthesis (1 min per 1000 base pairs):	68°C, 6’

Following the temperature cycling, the product was treated with Dpn I (10 U, 37°C, for 2 h). The Dpn I endonuclease (target sequence: 5´-Gm^{6}ATC-3´) is specific for methylated and hemimethylated DNA and is used to digest the parental DNA template and to select for mutation-containing synthesized DNA. 2 µl of the mixture were used to transform XL1-Blue or Top10 chemically competent cells. Plasmid DNA was isolated using QIAprep Spin Miniprep Kit (Qiagen) and was subjected to verification by automated sequencing.

2.2.2.11 Agarose gel electrophoresis of DNA

For verification of the presence and length of PCR or restriction digestion products, agarose gel electrophoresis was performed. For this purpose 1% agarose in a TAE buffer supplemented with ethidium bromide was prepared. The DNA samples were mixed with the 6x sample buffer prior to loading. DNA samples were run along with the 100bp and/or 1kb DNA ladder (NEB or pEQ lab) at 100-120 V DC. Results were visualized using UV illumination.

2.2.3 Transformation of *E. coli*

2.2.3.1 Preparation of chemically competent cells

A single colony of overnight grown bacteria from a LB agar plate was inoculated into 100 ml of LB media in a 500 ml flask. The culture was incubated at 37°C with vigorous agitation, monitoring the growth of cells. Cells were grown to an OD_{600} of ~0.6. The bacterial culture was transferred to sterile, disposable, ice-cold 50 ml polypropylene tubes and cooled down to 4°C on ice for 10 min. Cells were recovered by centrifugation at 3000 g for 10 min at 4°C. Supernatant media was decanted and tubes were kept in an inverted position on a pad of paper towel for 1 min to allow the last traces of media to drain away. Pellets were resuspended by gentle vortexing in 30 ml of the ice-cold MgCl2 solution. Again, cells were recovered by centrifugation at 3000 g for 10 min at 4°C. Supernatant solution was decanted and tubes were kept in an inverted position on a pad of a paper towel for 1 min to allow the last traces of solution to drain away. The pellet of the cells was recovered by gentle vortexing in 2 ml of ice-cold 0.1 M CaCl2

containing 15% glycerol, for each 50 ml of original culture. After this cells were dispensed into aliquots of 50 µl, flash frozen in liquid nitrogen and stored at -70°C.

2.2.3.2 Transformation of chemically competent cells

A number of factors have been elucidated that produced an increase in transformation efficiency. Such factors include: prolonged incubation of bacteria with $CaCl_2$, addition of multiple cations, such as Mg^{2+}or Ca^{2+} into the transformation mixture and treatment of bacteria with dimethyl sulfoxide (DMSO), polyethylene glycol, hexaminecobalt, and dithiothreitol in the presence of both monovalent and divalent cations (Chung et al., 1989). After incubation with DNA, in order to make the cells retain the plasmid and to be certain that they survive, the cells were heat shocked for several seconds to induce heat shock genes, which aid in survival and recovery. The cells were than incubated at 37°C without selective pressure; sufficient time was given for expression of antibiotic resistance genes. Plating on selective media enabled recovery of those cells that actually received the DNA.

Typically 3 µl of a ligation mix or ca. 50 µg of plasmid DNA was added to 50 µl of chemically competent cells. The mixture was incubated on ice for 30 min followed by a heat shock of 45 s at 42°C, 2 min cooling on ice, and the addition of a 250 µl of glucose and magnesium containing medium. After 1 h of incubation at 37°C, 20-50 µl of the mixture was spread out on LB agar plates (supplemented with selective antibiotics) and incubated overnight at 37°C.

2.2.4 Electroporation of *P. pastoris*

2.2.4.1 Preparation of competent cells

First a suitable *P. pastoris* strain was grown in 5 ml YPD in a 50 ml sterile falcon tube at 30°C overnight. Next 500 ml of fresh medium in a 2 liter flask was inoculated with 0.5 ml of the overnight culture. The culture was carried out overnight again to an OD_{600} = 1.3-1.5. Afterwards the cells were centrifuged at 1500 x g for 5 minutes at +4°C. The obtained pellet was resuspended with 500

ml of ice-cold (0℃), sterile water, centrifuged again and then resuspended with 250 ml of ice-cold (0℃), sterile water and centrifuged. Finally the pellet was resuspend in 1 ml of ice-cold (0℃) 1 M sorbitol for a final volume of approximately 1.5 ml. The cells were kept on ice and used the same day.

2.2.4.2 Transformation by electroporation

80 µl of the prepared competent cells were mixed with 50µg of DNA (in 10 µl sterile water) and transferred into an ice-cold (0℃) 0.2 cm electroporation cuvette. In order to make the cells retain the plasmid the cuvette with the cells was incubated on ice for 5 minutes. Then the cells were pulsed according to the manufacturer's instructions for yeast (*Saccharomyces cerevisiae*). Immediately after electroporation 1 ml of ice-cold 1 M sorbitol was added to the cuvette and the content was transferred to a sterile 15 ml tube and incubated at 30℃ without shaking for 1 to 2 hours. Next, plating on selective media with 100 µg/ml Zeocin enabled recovery of those cells that actually received the DNA. The plates were incubated 3-10 days at 30℃ until colonies formed.

2.2.5 Protein expression

The aim of expression is to get the maximum amount of a soluble and folded protein or in case when the recombinant protein is expressed only as insoluble inclusion bodies the maximum yield is requested. A number of parameters are significant to get the maximum yields of the protein, which include a type of culture media, temperature of culture growth, induction duration, induction OD, concentration of inducer and cell types.

The basic items for the choice of the appropriate expression system were covered in detail in chapter 2.2.1. In case of *E. coli* expression also the choice of the suitable host strain is important. The most popular *E. coli* system, is based on IPTG-induced expression under T7 promoter and BL21(DE3) strains are optimized for such protein expression. The BL21 strain is naturally deficient in the OmpT and Lon proteases, resulting in a higher yield of intact recombinant proteins. The suffix DE3 indicates that the host is a lysogen of *DE3, carrying a

chromosomal copy of the T7 RNA polymerase gene under the control of the IPTG-inducible lacUV5 promoter.

In this work the mainly used strain BL21 Star(DE3) from Invitrogen enhances additionally protein expression and yield by stabilizing mRNA. The cells carry a mutation in the gene encoding RNaseE, thereby eliminating a major source of mRNA degradation. As a result, more mRNA is available for protein translation. In some cases high-level expression of exogenous proteins may be toxic to the cells, some strains also carry a plasmid that encodes for T7 lysozyme (pLysS), a natural inhibitor of T7 RNA polymerase. These plasmids prevent leaky expression in uninduced cells by reducing basal expression of the polymerase. Even more refined control of IPTG-induced protein expression is offered by Tuner strain from Novagen. The strain is a Lac permease (lacY) deletion mutant that cannot utilize lactose and prevents Lac permease-mediated active transport of IPTG. Since IPTG can also independently enter *E. coli*, this strain allows for a truly concentration-dependent protein induction simply by adjusting culture concentrations of IPTG. Heterologous eukaryotic recombinant genes containing rare codons can be more effectively expressed in Stratagene's CodonPlus and Rosetta starins from Novagen. These competent cells contain extra copies of the genes that encode tRNAs for codons that are rarely used in *E. coli* and thereby guard against codon bias.

2.2.5.1 Standard *E. coli* expression in LB medium

10 ml LB were inoculated with a fresh single bacterial colony and incubated overnight at 37°C with vigorous shaking (200 rpm) in a 50 ml falcon tube. These overnight cultures were used as an inoculum for a pre-warmed 1 l LB medium, supplemented with an appropriate antibiotic. Cultures were incubated at 37°C with vigorous shaking (160 rpm) until the OD_{600} reached 0.8 - 0.9. Then the cells were induced by addition of 0.2 - 1 mM IPTG (final concentration) and grown with shaking until the expected OD was reached. For expression performed at 15-20°C cultures were grown after induction for about 12 h, for expression at 37°C between 3 and 4 h. Cells were harvested by centrifugation at 4000 rpm in a

Beckman centrifuge for 20 min at 4°C. The cell pellet after the expression was used directly for protein purification or stored at -80°C for longer time or -20°C if to be used in 1-2 days.

2.2.5.2 *E. coli* expression in the minimal medium (MM)

First 5 ml LB were inoculated with a single colony and shaken (200 rpm) 8-10 h in a 15 ml falcon tube at 37°C. Then 100 ml MM were inoculated with 100 µl of the initial culture and shaken (150rpm) overnight in a 250 ml flask at 37°C. The overnight culture was used to inoculate 1 l MM. Cultures were incubated at 37°C with vigorous shaking (160 rpm) and induced with IPTG (final concentration 0.2 - 1.0 mM) at OD_{600} = 0.7 -0.8. After an overnight incubation at 20°C with vigorous shaking (160 rpm), the cells were harvested by centrifugation at 4000 rpm at 4°C in a Beckman centrifuge for 20 min. The cell pellet was used directly for the protein purification or stored at -80°C for longer time or -20°C if to be used in 1-2 days.

2.2.5.2 *P. pastoris* expression

10 ml YP were inoculated with a fresh single yeast colony and incubated overnight at 30°C with vigorous shaking (200 rpm) in a 50 ml falcon tube. These overnight cultures were used as an inoculum for a pre-warmed 1 l YP medium, supplemented with an appropriate antibiotic. Cultures were incubated at 30°C with vigorous shaking (160 rpm) 2-3 days. Then the cells were centrifuged at 1500 x g for 5 minutes at +4°C. Obtained pellet was resuspended with 1 l of ice-cold (0°C), sterile water, centrifuged again and then resuspended in 1 l of BMMH medium. The culture was carried on 72 h. Medium containing expressed protein was separated from cells by centrifugation. In case of intracellular expression cells were harvested by centrifugation at 4000 rpm in a Beckman centrifuge for 20 min at 4°C. The cell pellet after the expression was used directly for protein purification or stored at -80°C for longer time or -20°C if to be used in 1-2 days.

2.2.6 Protein chemistry methods & techniques

Optimization of the conditions is important for purifying a protein. Sometimes the recombinant protein is expressed only as insoluble inclusion bodies. The protein might be solubilized and then refolded back to the native form by employing various refolding strategies.

In this work both methods of protein purification were applied to obtain soluble product. A number of parameters were checked to get the maximum yields of the protein, which include choice of appropriate strategies and their sequence, optimization of buffers used for purification and obtained protein concentration.

2.2.6.1 Protein purification under native conditions

Bacterial pellets were resuspended in the suitable binding buffer (20 ml of buffer for 1 l colony). The cell suspension was sonicated and centrifuged at 20000 rpm for 30 min at 4°C. The obtained supernatant was incubated with a Ni-NTA slurry (Qiagen) equilibrated previously in the binding buffer, for 2 h at 4 °C with gentle agitation. Next, the mixture was loaded onto an empty column and washed with washing buffer. The protein was eluted with the elution buffer. All buffers are described in section 2.1.12.

2.2.6.2 Protein purification under denaturating conditions

In the case of insoluble proteins accumulated in inclusion bodies, solubilization and refolding of the proteins into native states must be carried out. Solubilization must result in a monomolecular dispersion and the minimum intra- or inter-chain interactions. Bacterial pellets were first resuspended in PBS buffer (in order to remove soluble bacterial proteins), sonicated and centrifuged at 20000 rpm for 30 min at 16°C. The obtained supernatant was removed. Solubilization was achieved by approximately 12 h incubation, accompanied by vigorous stirring of obtained inclusion bodies in solutions containing 6 M guanidinium hydrochloride (binding buffer A), under strong reducing conditions (20 mM β-ME). Insoluble particles were removed and samples were clarified by

centrifugation at 25000 rpm for 30 min at 16°C. The obtained supernatant was incubated with a Ni-NTA slurry (Qiagen), equilibrated previously in buffer A, for 1 h at 37°C with gentle agitation. Next, the mixture was loaded onto an empty column and washed with buffer B. The protein was eluted with buffer C (section 2.1.12).

2.2.6.3 Sonication

Pulsed mode of operation was applied (output control 8, 60% duty cycle) and sonication was carried out on ice, in 5 steps of 3 min each, with 5 min intervals between steps, to avoid overheating of the sample.

2.2.6.4 Immobilized metal affinity chromatography (IMAC)

Presence of the 6-His-Tag in all the expressed proteins and complexes enabled the use of nickel affinity chromatography under reducing conditions. The pH of a supernatant prepared as described above was adjusted to 8.0 with 1 M NaOH (as optimal for binding to the Ni-NTA resin). A Ni-NTA slurry was added and binding was performed for 1-2 h with agitation. The ratio of the Ni-NTA matrix to the amount of the His-tagged protein is crucial for the purity of the protein. Usage of a small amount of resin and performing a stepwise elution allows obtaining a pure, concentrated protein in a short time.

2.2.6.5 Refolding

After an initial IMAC purification performed under denaturing conditions the refolding of the protein was carried out. Refolding is initiated by reducing the concentration of a denaturant. There are no strictly set rules to achieve the best conditions for refolding proteins from inclusion bodies. This can be performed by one- or multi-step dialyses, buffer exchange by gel filtration, applying a gradient of a refolding buffer when, the protein of interest is immobilized on the IMAC matrix (solid phase refolding), or by dilution (Tsumoto et al., 2003). During this work refolding was performed by both rapid dilution and dialysis. Many variables can be tried, like addition of arginine or glutamic acid, reduced/oxidized

glutathione ratio, urea, temperature, salt concentration, protease inhibitor, glycerol etc. The volume of the refolding buffer was adjusted to the total amount of protein to maintain its final concentration at a range of 0.1 – 0.2 mg/ml. The volume ratio of the denaturant containing sample to the buffer was kept at 1:50, resulting in the final Gua-HCl concentration of less than 0.15 M. For the rapid dilution technique, the protein sample was diluted in the refolding buffer (1:100) in ~150 µl aliquots, with 5-10min intervals between aliquots. Due to the gradual dilution, the accumulation of high concentration of folding intermediates was avoided, which could otherwise lead to aggregation. The refolding mixture was left with stirring at 4°C for 12 h. For the dialysis method, 2 l of the dialysis buffer were prepared. The protein was diluted to the concentration 0.5 mg/ml and dialyzed overnight.

2.2.6.6 Ion exchange chromatography

A high-resolution chromatographic medium, MonoQ, could be used at a second purification step. Monodispersive, 5-µm bed support particles and a large protein binding capacity of the 8-ml column, ranging up to 20 mg/ml, assured excellent resolution of absorbed proteins. Bound proteins were fractionated with a linear NaCl gradient, the length of which was chosen individually for each construct on the basis of pilot experiments.

2.2.6.7 Protease digestion

In case of differently tagged (mostly His-tagged) proteins the tag was removed prior to crystallization trials by a suitable protease digestion. The reaction was performed according to the manufacturer manual. Optionally if the removed tag was of similar size to the protein, the protein was subsequently purified over NiNTA to remove His–tags or undigested proteins. After digestion all proteins were purified on the gel filtration HiLoad 26/60 Superdex S75 or S200pg Amersham Pharmacia column. The His-tag removal for all constructs was controlled by using SDS–PAGE and Western–blots.

2.2.6.8 Gel filtration chromatography

This final step of purification enabled separation of the proteins from digested tag peptides, and the restriction protease used for tag cleavage. Complete exchange of the buffer was also possible, removing any low molecular weight substances, which could interfere with NMR measurements (such as Tris, EDTA, protease inhibitors, glycerol etc.) The column used, a 120-ml Superdex 75, is characterized by good resolution in the range 3-60 kDa. The 120-ml Superdex 200 allows excellent separation of proteins in the range from 20 to 250 kDa. To make the best of its capability, low flow rates (0.75-1.25 ml/min) were used and samples not larger than 6 ml were loaded on the column.

2.2.7 Handling and storing of the proteins

All the operations involving the proteins at either level of purification were performed on ice or at 4℃, unless denaturating conditions were used. 2-mercaptoethanol was added to buffers during purification of the proteins containing Cys - to prevent disulfides bond formation. NaN_3 was added to all solutions at concentration 0.05%. The function of NaN_3 was to prevent bacterial and fungal growth in solutions stored for prolonged time. Purified proteins were stored at +4℃ or at -80℃ to further analyses.

2.2.8 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

The SDS polyacrylamide gel electrophoresis was carried out at various stages of purification to check the purity and identity of the eluted proteins. For all of the expressed proteins, tricine gels were applied (Schagger and von Jagow, 1987). The protein samples were prepared by mixing 20 µl of protein solution with 5 µl of sample buffer (SB) followed by 5 min incubation at 100 ℃. Due to rapid precipitation of SDS in contact with guanidine, the samples (after Ni-NTA chromatography under denaturing conditions) were prepared as follows: 20 µl of the protein solution in a denaturing buffer was diluted with 400 µl 20% trichloroacetic acid (TCA). The sample was incubated for 5 min on ice followed by centrifugation for 5 min at 20 000 x g. Supernatant was discarded by suction,

the precipitated protein pellet was washed once by vortexing with 400 µl ethanol. After centrifugation and ethanol removal, the protein pellet was resuspended in 20 µl of 2x SB and the sample was boiled for 5 min.

2.2.9 Visualization of separated proteins

For visualization of the protein bands, the gels were stained in a Coomassie-blue solution. The background was cleared by incubation of the gel in a destaining solution. Both processes were greatly accelerated by brief heating with microwaves of the gel submerged in an appropriate solution.

2.2.10 Western blot

Western blot is a functional assay to confirm the immunological identity of the expressed proteins or identify the protein out of a number of proteins. It is helpful also to determine the presence of any degradation products or isoforms of the purified proteins. The semi-dry Western blot was applied. For the blot the nitrocellulose membrane and six Watmann papers of the size of the SDS-PAGE gel were cut and soaked in the transfer buffer. Watmann paper, the protein separated previously by SDS-PAGE (gel), and nitrocellulose membrane were arranged in the following order over the electroblot: three wet Watmann papers, wet nitrocellulose paper, and SDS-PAGE gel followed by three wet Watmann papers. The apparatus was closed and run (transfer) at constant voltage of 100 V for 1-1:30 h. After the transfer, the nitrocellulose membrane was taken and kept in the blocking solution (5% milk powder solution) for 2 h with gentle agitation. After blocking, the membrane was washed three times with the wash buffer and incubated for 1.5 h at room temperature with the primary antibody solution (the procedure can be stopped at this point by keeping the blot in the antibody solution at 4°C overnight). The membrane was washed with the wash buffer and incubated in the secondary antibody solution for 1.5 h at room temperature. After this the membrane was washed three times with the wash buffer and the blot was developed by incubating it in the substrate (BCIP) solution for 10 min.

2.2.11 Determination of protein concentration

The concentration of proteins in solution was also estimated by means of the Bradford colorimetric assay (Bradford, 1976). 5 µl of the protein sample was added to 1 ml (10 x diluted stock) of Bradford reagent (BioRad) in a plastic cuvette. After gentle mixing, A_{595} was measured and converted to the protein concentration on the basis of a calibration curve prepared for known concentrations of BSA.

Determination of protein concentration was performed spectrophotometrically. Absorption at 280 nm was measured and converted to a protein concentration on the basis of theoretical extinction coefficients. It has been shown that it is possible to estimate the molar extinction coefficient $E_\lambda(Prot)$ of a protein from knowledge of its amino acid composition (Gill and Hippel, 1989). From the molar extinction coefficient of tyrosine, tryptophan and cysteine (cysteine residues do not absorb appreciably at wavelenghts >260 nm, while cysteine does) at a given wavelength λ the extinction coefficient of a protein can be computed using the equation:

$$E_\lambda(Prot) = Numb(Y)xExt_\lambda(Y) + Numb(W)xExt_\lambda(W) + Numb(C)xExt_\lambda(C)$$

Protein concentration (C_p) can be calculated using the following formula:

$$A_\lambda\ (Prot) = E_\lambda(Prot) \times C_p(Prot) \times (\text{cuvette path length in cm})$$

2.2.12 Mass spectrometry

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

2.2.13 NMR spectroscopy

All NMR experiments were carried out at 300 K on a Bruker DRX 600 spectrometer equipped with a triple resonance, triple gradient 5 mm cryoprobehead. The spectra were analyzed with the program Sparky (Goddard,

and Kneller, University of California, San Francisco). The samples contained typically 0.1-0.5 mM protein preferably in a PBS buffer supplemented with 10% ^2H_2O. All 1D ^1H NMR spectra were recorded with a time domain of 32 K complex points and a sweep-width of 10,000 Hz. The 2D ^1H-^{15}N-HSQC spectra were recorded with a time domain of 1K complex data points with 128 complex increments with a sweep width of 8 kHz in the ^1H dimension and 2 kHz in the ^{15}N dimension.

2.2.14 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a thermodynamic technique for monitoring any chemical reaction initiated by the addition of a binding component, and has become the method of choice for characterizing biomolecular interactions. When substances bind, heat is either generated or absorbed. Measurement of this heat allows accurate determination of dissociation constants (K_D), reaction stoichiometry (n), enthalpy (ΔH) and entropy (ΔS), thereby providing a complete thermodynamic profile of the molecular interaction in a single experiment. All ITC experiments were carried out according to references provided by the manufacturer. Investigated proteins were used at 0.03 - 0.04 mM in PBS and titrated from a 300 µl syringe into a sample chamber holding 1.43 ml of 0.3 - 0.4 mM of a respective binding partner. All solutions were degassed prior to measurements. Heat generated by protein dilution was determined in separate experiments by injecting a protein solution into a PBS filled sample chamber. All data were corrected for the heat of protein dilution. Data were fitted using χ^2 minimization for a model, assuming a single set of binding sites to calculate the binding affinity K_D. All steps of the data analysis were performed using the ORIGIN (V5.0) software provided by the manufacturer. The details of the experimental and injection parameters are described below:

total number of injections:	30
volume of a single injection [µl]:	10
duration of an injection [s]:	24

intervals between injections [s]:	600
filter period [s]:	2
equilibrium cell temperature [°C]:	20
initial delay [s]:	60
reference power [µCal/s]:	15
stirring speed [RPM]:	270

2.2.15 Crystallization trials

2.2.15.1 Protein preparation

Both, sitting drop and hanging drop vapor diffusion techniques were used for crystallization. The final purification of proteins was against 10 mM Tris pH 8, 100 mM NaCl, 0.05% NaN3 for Mad2 constructs and complexes and 10 mM Hepess pH 7.4, 15 mM NaCl, 0.05% NaN3 for CDK2. Collected fractions were pooled and concentrated until the necessary protein concentration was achieved (5-30 mg/ml). Subsequently, the samples were centrifuged to clear out small particles and precipitates. The crystallization procedure was performed as quickly as possible (usually directly after concentration of the proteins). To determine preliminary crystallization conditions the Hampton Research Crystal Screens were used. 0.1 µl of a protein sample and 0.1 µl of the crystallization buffer were mixed in 96-well plates; the reservoirs were filled with 100 µl of buffer. Crystallization plates were set up at two temperatures: 4°C and room temperature. After crystals appeared the crystallization was performed on a larger scale in the best conditions of the initial screen. 1 µl of the protein sample and 1 µl of the crystallization buffer were mixed in 24-well plates. Reservoirs were filled with 500 µl of crystallization buffer.

2.2.15.2 Data collection and structure analysis

The data for all crystals were collected from shock frozen crystals at a rotating anode laboratory source. Prior to freezing, the crystals were soaked for 30 s in a drop of a reservoir solution containing 15% v/v ethylene glycol or 15% glycerol as cryoprotectant. The high-resolution datasets were collected on the

MPG/GBF beamline BW6 at DESY, Hamburg, Germany. Collected data were integrated, scaled and merged by XDS and XSCALE programs (Kabsch, 1993). The structure was determined by molecular replacement using the Molrep program from the CCP4 suite (CCP4, 1994).

2.2.17 Thermal shift assay

Protein melting point analysis was performed on a LightCycler 480 RT-PCR machine (Roche Applied Science). In a total volume of 20 µl, the 30 µM protein was combined with SYPRO Orange (Molecular Probes/Invitrogen) at 2.5 x final concentration. The time/temperature control of the PCR machine was programmed to perform 120 steps of 1 min each and raising the temperature at each step by 0.5°C, from 37°C to 95°C. The fluorescence emitted after excitation at 480 nm was recorded at each step.

3 IGF1-R

3.1 Biological background

The communication between the cells is of great significance for their successful and efficient functioning. Especially in multicellular organisms cells have to posses the ability to communicate with one another in complex ways to be able to govern their own behavior for the benefit of the organism as a whole.

The communication mechanisms usually act through extracellular molecules, which are produced by cells to send a signal to their neighbors or to cells further away. These molecules bind to the appropriate cell-surface receptor and activate a signal cascade, the so called “signaling pathways”. They employ sophisticated systems of proteins, which include a variety of intracellular signaling proteins followed by target proteins. The differences between effector proteins that each cell contains enable them to respond to a particular subset of these signals in a specific way (for a comprehensive overview see Alberts et al., 2002).

The signals received from membrane receptors activated by various specific factors can modulate a gene expression pattern of the cell leading to differentiation. It can also decide fates of the mature cell, like for example, the rate of nucleic acid and protein synthesis or cell cycle progression. In a cell deprived of this steady stimulation suicidal pathways take over, leading to apoptosis.

3.1.1 The structure and function of the IGF system

The insulin-like growth factor (IGF) system is a highly conserved signaling pathway (Figure 3.1.1). It is composed of two IGF receptors, two IGF ligands, and six IGF high-affinity binding proteins (IGFBPs). IGF ligands, IGF-1 and IGF-2, bind to the insulin/IGF family of cell surface receptors while the high affinity IGFBPs modulate the availability and bioactivity of IGF-1 and -2. Almost all cell

types express IGF receptors and therefore respond to IGF stimulation and the IGF system plays a key role in the regulation of cell development and survival.

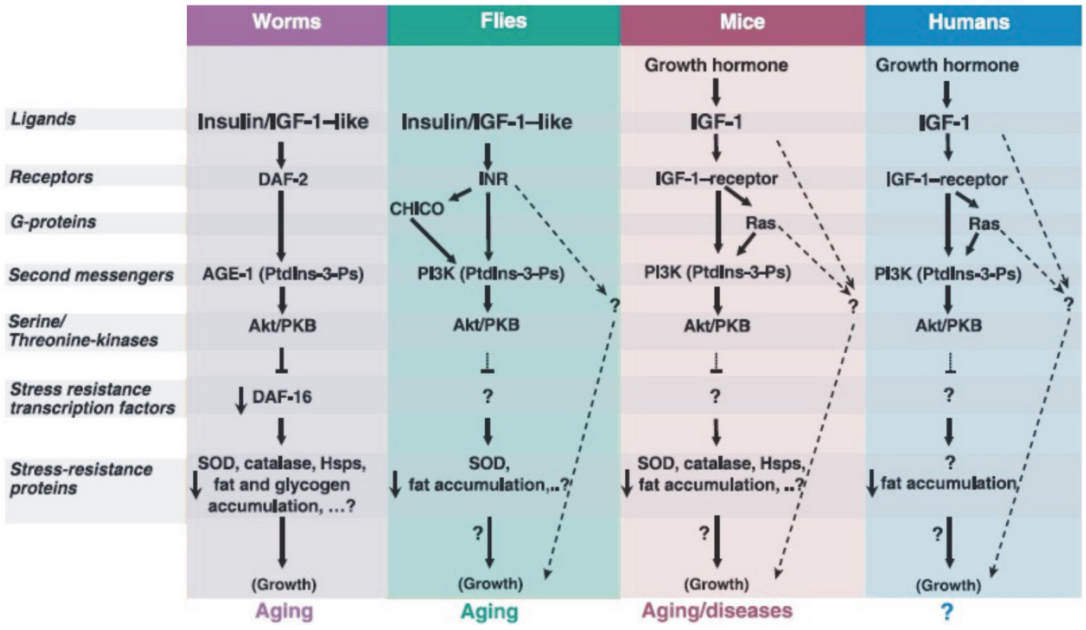

Figure 3.1.1. Conserved regulation of longevity. In worms, and flies, the partially conserved insulin/IGF-I–like pathways down-regulate antioxidant enzymes and heat shock proteins, reduce the accumulation of glycogen or fat, and increase growth and mortality. Figure adapted from Longo and Finch, (2003).

The IGF signaling pathway arose early in evolution, possibly as a regulator of cellular proliferation in relation to nutrient availability (Longo and Finch, 2003). This function is conserved in mammalians - a number of biological processes including cellular growth, proliferation, differentiation, survival against apoptosis and migration are under its control. It mediates the growth-promoting effect of growth hormone (Schlechter et al., 1986; Daughaday et al., 1999), promotes nerve cell survival and regeneration (Leventhal et al., 1999; Jung et al., 1998) and acts in several sites during the repair process (Suh et al., 1992). In the kidney, IGF-1 regulates the glomerular filtration rate and phosphate transport (Halloran and Spencer, 1988). This way the IGF system is involved in tissue formation and remodeling, bone growth, brain development, and regulation of metabolism.

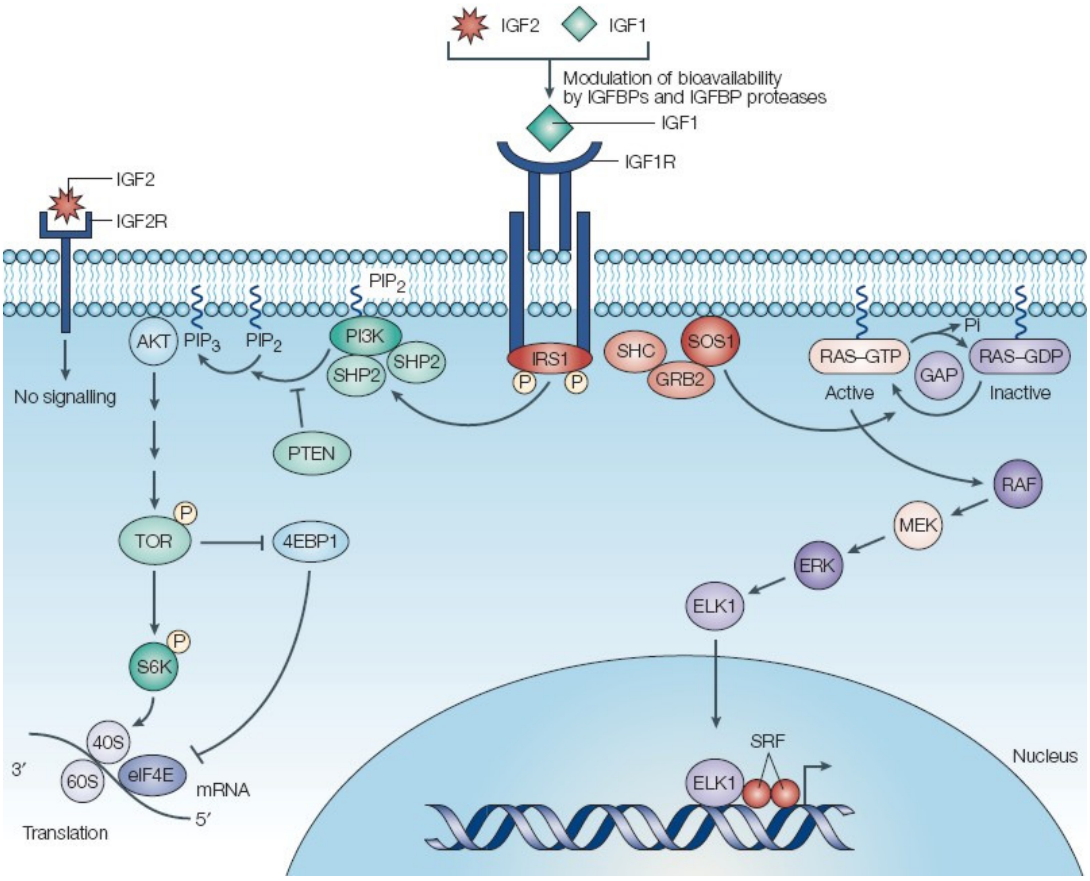

Figure 3.1.2. Overview of IGF-1R activation and downstream signaling. The IGF-1R binds either IGF-1 or IGF-2. Ligands can be delivered from remote sites of production through the circulation or be locally produced. IGFBPs and IGFBP proteases have key roles in regulating ligand bioavailability. The IGF-2R binds IGF-2, but has no tyrosine kinase domain and appears to act as a negative influence on proliferation by reducing the amount of IGF2 available for binding to IGF-1R. Ligand binding to IGF-1R, activates its tyrosine kinase activity, and this stimulates signalling through intracellular networks that regulate cell proliferation and cell survival. Key downstream networks include the PI3K–AKT–TOR system and the RAF–MAPK systems. Activation of these pathways stimulates proliferation and inhibits apoptosis. 4EBP1, eukaryotic translation initiation factor 4E binding protein 1; eIF4E, eukaryotic translation initiation factor 4E; ERK, extracellular signal-regulated kinase; GRB2, growth-factor-receptor-bound protein 2; IRS1, insulin-receptor substrate 1; MAPK, mitogen-activated protein kinase; MEK, mitogen-activated protein kinase kinase; PI3K, phosphatidylinositol 3-kinase; PIP, phosphatidylinositol; PTEN, phosphatase and tensin homologue; S6K, S6 kinase; SHC, SRC-homology-2-domain transforming protein; SHP2, phosphatidylinositol 3-kinase regulatory subunit; SRF, serum response factor; TOR, target of rapamycin. Figure reproduced from Pollak et al., (2004).

3.1.2 IGF-1 and IGF-2

Human IGF-1 and IGF-2 are evolutionarily conserved single-chain proteins. They display approximately 62% sequence similarity to each other. The mature IGF-1 is a basic polypeptide containing 70 amino acids while IGF-2 is slightly acidic and contains 67 amino acids. Both form three disulfide bonds (Cys6–Cys48; Cys18–Cys61; Cys47–Cys52 for IGF-1; Cys9–Cys47; Cys21–Cys60; Cys46–Cys51 for IGF-2) and are organized in four domains named A, B, C, and D. The IGF A and B domains are similar to insulin A and B chains, respectively, with about 50% sequence similarity (Figure 3.1.3). The C domain is analogous to the connecting peptide of proinsulin. The D region, C-terminal peptide shows no homology to insulins and proinsulins (Baxter et al., 1992).

Chain A:

```
IGF-1     42 GIVDECCFRSCDLRRLEMYCA 62
IGF-2     41 GIVEECCFRSCDLALLETYCA 61
SCI       A1 GIVEQCCTSICSLYQLENYCN 21
```

Chain B:

```
IGF-1      1 GPETLCGAELVDALQFVCGDRGFYFNKPT 29
IGF-2  1 AYRPSETLCGGELVDTLQFVCGDRGFYFSRPA 32
SCI        1 FVNQHLCGSHLVEALYLVCGERGFFYT-PK 29
```

Chain C:

```
IGF-1     30 GYGSSSRRAPQT 41
IGF-2     33 S--RVSRRSR 41
```

Chain D:

```
IGF-1     63 PLKPAKSA 70
IGF-2     62 T--PAKSE 67
```

Figure 3.1.3. Sequence alignment of IGFs and single-chain insulin (SCI). Residues for binding to IGF-1R are highlighted in red.

Several three-dimensional structures of IGFs by both NMR and X-ray crystallography have been determined (Cooke et al., 1991; Sato et al., 1993; Schaffer et a., 2003; Vajdos et al., 2001). The overall structure of IGF-1 and IGF-

2 within the A and B domains is similar to the crystal structure of insulin which has been first resolved in the late 1960s (Adams et al., 1969), and the NMR structure of proinsulin (Weiss et al., 1990). The major secondary structural elements of IGF-1 and IGF-2, as well as insulin, are α-helical. The A domain contains helix 2 (Ile43–Cys47 of IGF-1; Glu44–Phe48 of IGF-2) and helix 3 (Leu54–Glu58 of IGF-1; Ala54–Tyr59 of IGF-2) whereas the B domain is built of helix 1 (Gly8–Cys18 of IGF-1; Gly10–Val20 of IGF-2). The IGF C and D domains are unstructured and highly flexible in solution.

In the late 1980s the abbreviated form of IGF-1, so called DES(1-3)IGF-1, has been found in fetal and adult human brain (Carlsson–Skwirut et al., 1986; Sara et al., 1986; Humbel, 1990). The DES(1-3)IGF-1 is the product of differential processing of pro-IGF-1 lacking the first three residues at the amino terminus: Gly-Pro-Glu. The biological potency of this truncated form is 10 times higher than that of the full-length form and is explained to be due to reduced binding to IGF-binding proteins (Francis et al., 1988; Beck et al., 1993; Carlsson–Skwirut et al., 1989; Ballard et al., 1996).

In mammals, IGFs have characteristics of both a circulating hormone and a tissue growth factor. They are widely expressed during fetal and prenatal stages. In postnatal stages, hepatic production of IGF-1, under the regulation of growth hormone (GH), becomes the major source of circulating IGF-1 (Figure 3.1.4). GH has a dominant role in upregulating IGF-1 gene expression but its stimulatory influence is markedly reduced by malnutrition (Thissen, 1994). GH is produced by the pituitary gland under the regulation of the hypothalamic factors somatostatin (SMS) and growth-hormone-releasing hormone (GHRH). However, IGF-1 is also synthesized in many non-hepatic tissues where autocrine mechanisms of action are important (LeRoith and Helman, 2004). IGF-2 is also expressed both in the liver and in extrahepatic sites, but is not tightly regulated by GH (Pollak et al., 2004).

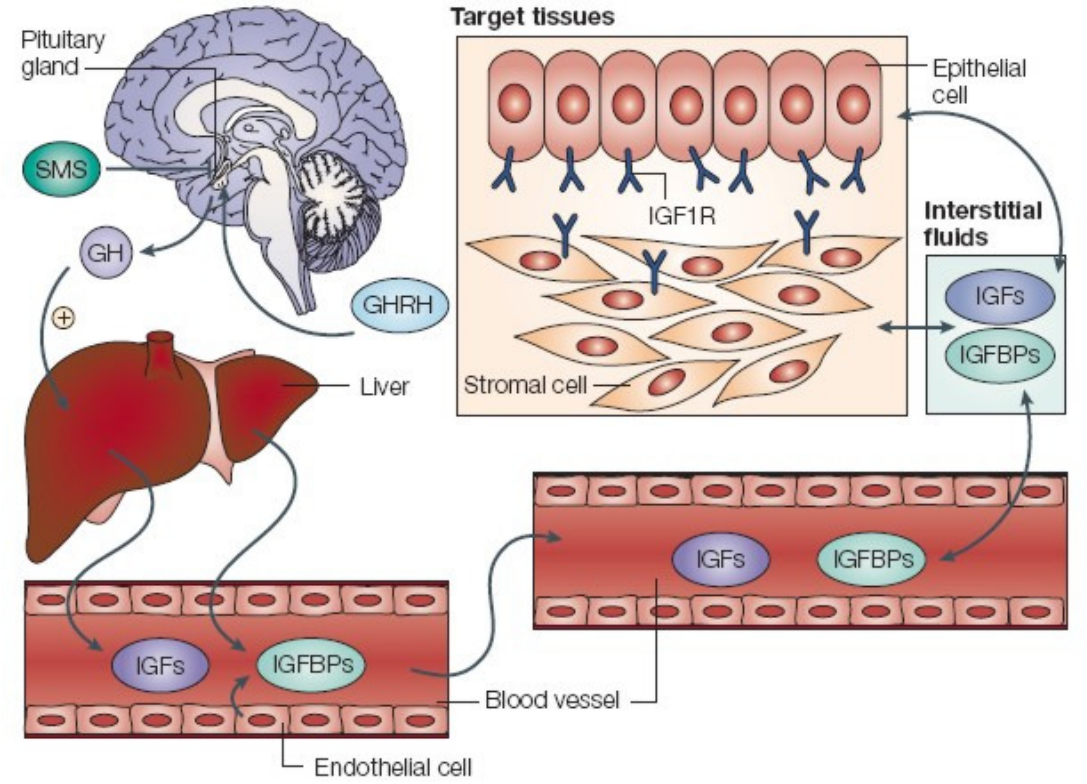

Figure 3.1.4. Regulation of circulating and tissue levels of IGFs. Most circulating IGFs are produced in the liver. Hepatic IGF1 production is subject to complex regulation by hormonal and nutritional factors. GH which is produced in the pituitary gland under control of the GHRH and SMS, is a key stimulator of IGF1 production. Various IGF-binding proteins (IGFBPs) are also produced in the liver. In IGF responsive tissues, the ligands IGF1 and IGF2 as well as IGFBPs can be delivered through the circulation from the liver (an 'endocrine' source), but IGFs and IGFBPs can also be locally produced through autocrine or paracrine mechanisms. These mechanisms often involve interactions between stromal- and epithelial-cell subpopulations. Figure reproduced from Pollak et al., (2004).

3.1.3 IGFBPs

The endocrine hormones like insulin or growth hormone are secreted from storage vesicles in an endocrine gland in response to releasing factors such as GH releasing factor or glucose. The IGFs are stored in the extracellular environment in a unique manner for a protein hormone, bound to specific IGFBPs. This way IGFBPs modulate bioavailability of IGFs in both the circulation and the cellular microenvironment (Firth and Baxter, 2002).

The IGFBP family consists of six multifunctional, homologous proteins (IGFBP-1 to -6) of a size between 216 to 289 residues. They are expected to have closely related IGF binding determinants and bind to IGFs with nanomolar affinities (Firth et al., 2001; Clemmons, 2001; Bunn and Fowlkes, 2003). All of them are synthesized as precursor forms with secretory signal peptides of between 20 and 39 amino acids, so the mature proteins are found extracellularly. Because of their sequence homology (Figure 3.1.5), IGFBPs are assumed to share a common overall fold and a common domain organization. Each IGFBP forms three distinct domains of approximately equal lengths: highly conserved cysteine-rich N- and C-domains and a central linker domain unique to each IGFBP species (Figure 3.1.6). This domain structure is highly conserved among this gene family and across species (Duan, 1997; Duan et al., 1999; Maures and Duan, 2002). Both the N- and C-domains participate in the binding to IGFs (Firth et al., 2001; Clemmons, 2001; Bach et al., 2005; Bunn and Fowlkes, 2003; Payet et al., 2003; Allan et al., 2006; Carrick et al., 2002; Kibbey et al., 2006; Fernandez-Tornero et al., 2005; Siwanowicz et al., 2005; Headey et al., 2004) , the specific roles of each of these domains in IGF binding have only been recently determined (Sitar et al., 2006).

The only one clearly identified function of the N-domains of IGFBPs is the IGF binding, whereas the carboxy-terminal domain may be responsible for preferences of IGFBPs for one species of IGF over the other. C-domains also interact with a large number of proteins and other biologically significant molecules and are involved in a wide range of functions (Bach et al., 2005). The central linker domain is the least conserved region and has never been mentioned as part of the IGF-binding site for any IGFBP. This domain is the site of post-translational modifications and specific proteolysis (Bunn and Fowlkes, 2003). Proteolysis is believed to produce low affinity N- and C-terminal fragments that cannot compete with IGF receptors for IGFs and thus it can be the predominant mechanism for IGF release from IGFBPs (Bunn and Fowlkes, 2003; Fernandez-Tornero et al., 2005).

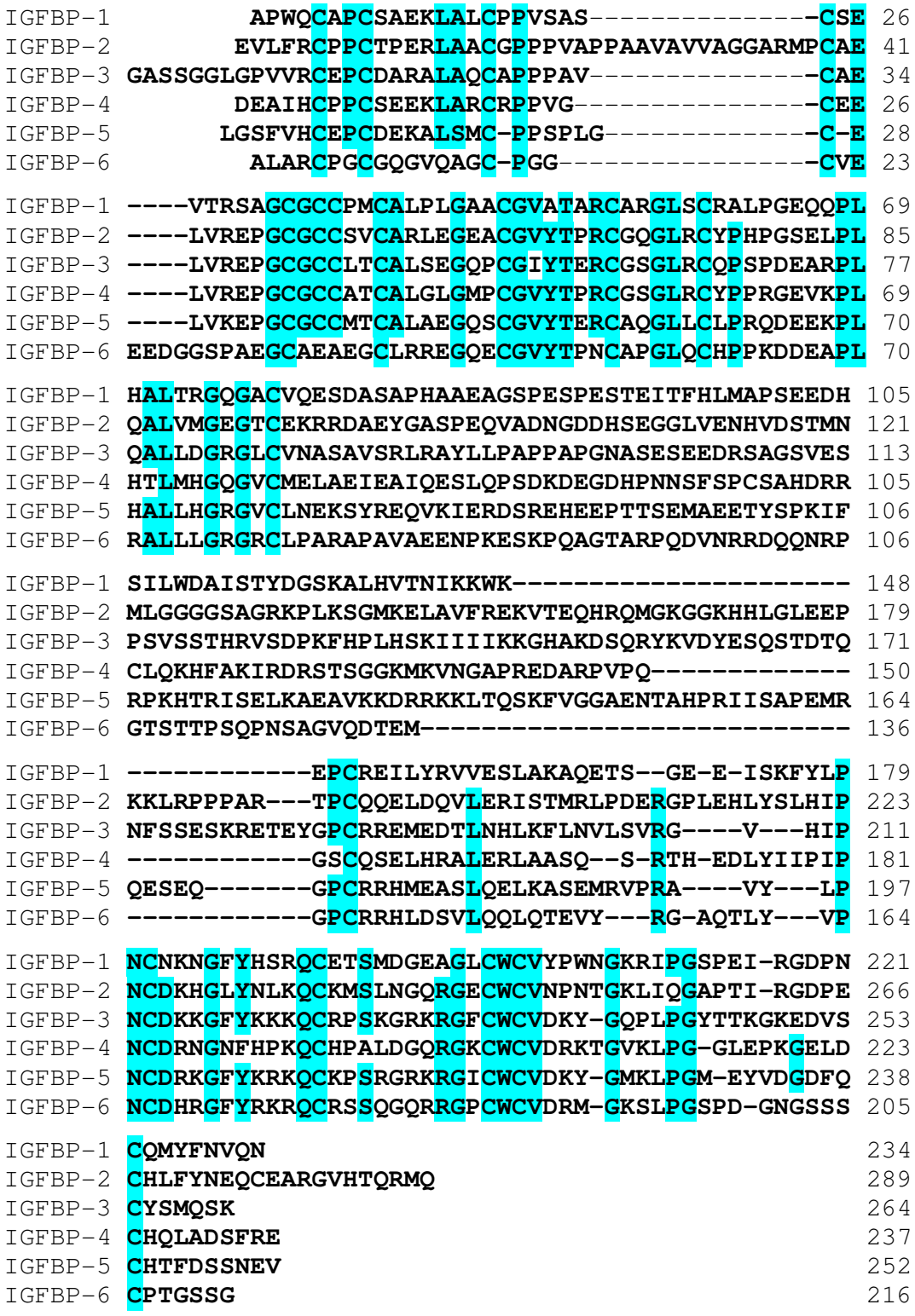

Figure 3.1.5. Sequence alignment of human IGFBP-1 to -6. Conserved residues are indicated by cyan shading.

However, recent studies indicate that the resulting N- and C-terminal fragments still can inhibit IGF activity and have functional properties that differ from those of the intact proteins (Firth and Baxter, 2002; Bach et al., 2005; Allan et al., 2006; Fernandez-Tornero et al., 2005)

Suppressing IGF-dependent actions was postulated to be the dominant function of IGFBPs. Though, recent data indicate additional complexities. Numerous IGF-independent effects mediated by IGFBPs C-domains have been described *in vitro* (Firth and Baxter, 2002; Ricort, 2004), and *in vivo* (Miyakoshi et al., 2001), including growth inhibition, promotion of apoptosis, and modulation of cell adhesion and migration (Perks et al., 1999; Pereira et al., 2004). Additionally, in certain physiological contexts, IGFBPs seem to increase rather than reduce the mitogenic activity of IGFs (Firth and Baxter, 2002). These actions might involve lengthening the half-life of IGFs or poorly characterized actions whereby IGFBPs deliver IGFs to IGF-1R (Pollak et al., 2004).

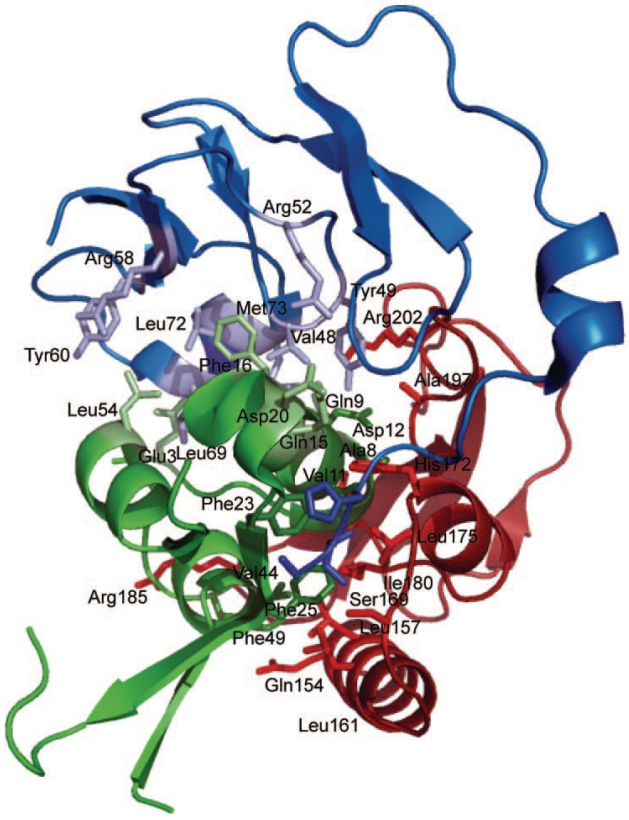

Figure 3.1.6. Overall structure and folding of the N-domain of the BP4(3–82)-IGF1-C-domain BP4(151–232) complex. The N-BP4 domain is in blue, IGF1 is in green, and C-BP4 is in red. The residues important for binding among protein components are indicated with numbers (Sitar et al., 2006).

3.1.4 IGF receptors

The IGFs, as well as insulin, interact with specific cell surface receptors (IGF-1R, IGF-2R, IR). Although the physiological functions of insulin and IGFs are very different, their receptor structures are similar. The insulin receptor is a phylogenetically ancient tyrosine kinase receptor found in organisms as primitive as cnidarians and insects. In higher organisms it is essential for glucose homeostasis, whereas the closely related IGF-1R is involved in normal growth and development (Adams et al., 2000; McKern et al., 2006). The IGF-1, type II tyrosine kinase, is a large transmembrane glycoprotein of 400-450 kDa, consisting of two α-subunits (135 kDa each) and two β-subunits (90 kDa each) (LeRoith et al., 1993; LeRoith et al., 1995; Rechler, 1985; Yamasaki et al., 1993). Disulphide bonds connect both α- and β-subunits to form a functional heterotetrameric receptor complex. In analogy with the insulin receptor, IGF-1 receptor subunits are encoded within a single 180 kDa polypeptide precursor that is glycosylated, dimerised and proteolytically processed to yield the mature $\alpha_2\beta_2$ form of the receptor. The α-subunit is entirely extracellular and contains the ligand-binding site. The β-subunit contains the hydrophobic transmembrane domain with a short extracellular region, and a tyrosine kinase domain in its cytoplasmic portion (Figure 3.1.7).

The structure of the first three domains of the extracellular portion of IGF-1R (leucine-rich repeat (L1) domain, a cysteine-rich (CR) region and a second leucine-rich repeat (L2) domain, residues 1-462) has been determined to a 2.6 Å resolution. The L domains each adopt a compact shape consisting of a single-stranded right-handed β-helix. The CR region is composed of eight disulphide-bonded modules, seven of which form a rod-shaped domain (Garrett et al., 1998). The first there domains of the IGF-1R differ structurally from the IR in the regions governing ligand specificity (Lou et al., 2006). The crystal structure of the IR ectodomain has been resolved recently (McKern et al., 2006). The structure reveals the domain arrangement in the disulphide-linked ectodomain dimer, showing that the insulin receptor adopts a folded-over conformation that places the ligand-binding regions in a juxtaposition. As well as for IGF-1R, each IR

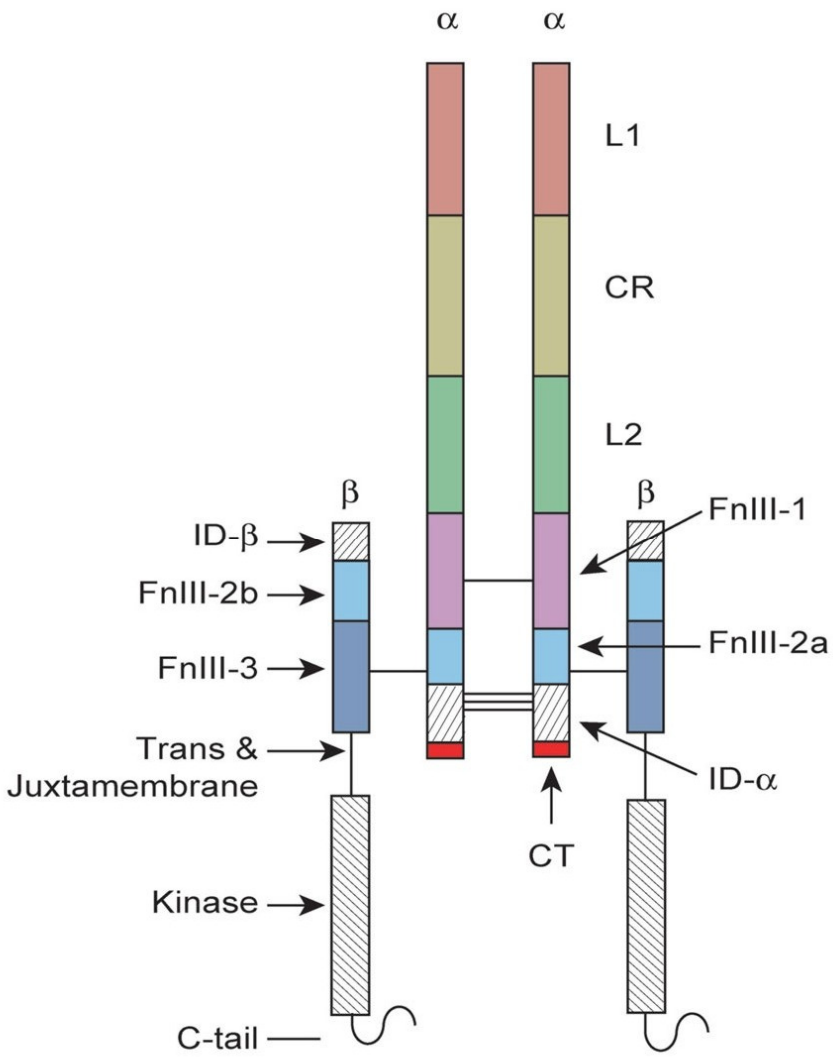

Figure 3.1.7. Schematic diagram of the IR homodimer Figure reproduced from McKern et al., (2006).

contains L1, CR and L2 domains, followed by three fibronectin type III domains (FnIII-1 to FnIII-3). FnIII-2 contains an insert domain (ID) of 120 residues, within which the α-β cleavage site is located. Each α-β monomer has an inverted 'V' layout with respect to the cell membrane. One leg of the 'V' is formed by the L1, CR and L2 domains; the other is formed by an extended linear arrangement of the three FnIII domains. In the ectodomain homodimer, the L2 domain of the first

monomer contacts the FnIII-1 domain of the second monomer at the apex of the inverted 'V', whereas the L1 domain of the first monomer contacts the FnIII-2 domain of the second at approximately the midpoint of one of the two legs of the inverted 'V'. This arrangement which is significantly different from previous models (De Meyts and Whittaker, 2002), shows that the L1 domains are on opposite sides of the dimer, too far apart to allow insulin to bind both L1 domains simultaneously (Figure 3.1.8).

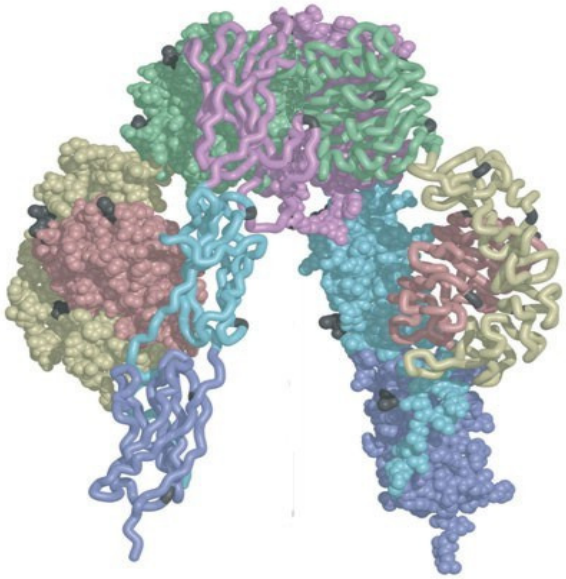

Figure 3.1.8. The IR ectodomain homodimer, showing the juxtaposition of domains. One monomer is shown in tube representation, the other in atomic sphere representation. The L1 domain is in brown; CR in yellow; L2 in green; FnIII-1 in magenta; FnIII-2 in cyan; FnIII-3 in blue. The potential N-linked glycosylation sites are shown in black. Figure reproduced from McKern et al., (2006).

The IGF-1 binding to the extracellular α-subunit of IGF-1R effects autophosphorylation of three tyrosines in the activation loop of the tyrosine kinase domain in the cytoplasmic portion of the β-subunit, which results in amplification of tyrosine kinase activity and further autophosphorylation of additional tyrosine residues. These phosphotyrosine-containing motifs are binding sites for adaptor and effector molecules in receptor signaling pathways, including insulin receptor substrates and SHC (Figure 3.1.2), which are

subsequently phosphorylated on their tyrosines (White and Kahn, 1989; Kim et al., 1998). The insulin receptor substrates (IRSs) are known as “docking” proteins and constitute a family of four structurally related adaptor proteins that can link the IGF-1 receptor to downstream signal transduction mediators regulating cellular growth. IRS 1 is the most extensively studied, and has multiple tyrosines, which associates with SH2 domain-containing proteins including the growth factor receptor bound-2 protein (Grb2) and the p85 regulatory subunit of phosphoinositol-3 (PI-3) kinase, p110. Phosphorylation of IRS 1 and 2 leads to activation of two downstream signaling cascades: the mitogen-activated protein kinase (MAPK) and the phosphatidylinositol 3-kinase (P3K) cascades (Baserga et al., 2003; Nakae et al., 2001; Jones and Clemmons, 1995).

3.1.5 The IGF system and cancer

A malfunctioning IGF signaling pathway is involved in many common diseases including atherosclerosis and diabetic complications. Reduced circulating IGF-1 levels are associated with type I diabetes, and the IGF-1 treatment improves glucose and protein metabolism and attenuates diabetic cardiomyophaty (Carrol et al., 2000; Norby et al., 2002). Recent studies suggest that insulin/IGF signaling is required for male sex determination: XY mice deficient in the insulin receptor, IGF-1R receptor and IR-related receptors (IRR) have a completely female phenotype (Nef et al., 2003). In humans, a homozygous partial deletion of the IGF-1 gene is associated with mental retardation and sensorineural deafness, in addition to fetal and postnatal growth retardation (Woods et al., 1996). In the central nervous system, expression of IGF-1 and the IGF-1R are induced by brain injury, and exogenous administration of IGF-1 after injury ameliorates the damage (Guan et al., 2003). Accumulation of β-amyloids in the brain during aging is associated with decreased levels of IGF-1 in serum. IGF-1 treatment delays progression of amyotrophic lateral sclerosis (ALS) in a mouse model (Kaspar et al., 2003).

Though, most significant is the association of the IGF system with cancer. Converging results from epidemiological studies and *in vivo* carcinogenesis

models indicate that increased risks of breast, prostate, colorectal, and lung carcinomas are linked with increasing serum concentrations of IGF-1 (LeRoith and Roberts, 2003; Furstenberger and Senn, 2002). Several model systems have provided evidence that proliferation and metastasis of tumor cells are increased by IGF-1R activation, either in relation to higher levels of circulating IGF1 in the host or to autocrine production of ligands by neoplastic cells (Kahandwala et al., 2000). These findings were confirmed in animal models, where reduced circulating IGF-1 levels result in significant reductions in cancer development, growth, and metastases (Wu et al., 2003). On the other hand the neoplastic progression, particularly in prostate cancer, is associated with overexpression of IGF-1R (Hellawell et al., 2002). Moreover, the increased expression of the *IGF2* gene observed in colon cancers indicates an important role for IGF2 expression in cancer progression (Zhang et al., 1997). Also downstream elements of IGF-1R signaling are implicated in neoplasia. Signals from IGF-1R can become exaggerated or inappropriate because of the molecular pathology (Pollak et al., 2004). A typical example is the loss of function of the tumor-suppressor gene *PTEN* which encodes a phosphatase that attenuates signals originating at tyrosine kinase receptors (Figure 3.1.2).

The concept of receptor targeting is now well established in anticancer therapeutics. Recently, success in the clinical studies of compounds that target tyrosine kinase receptors, such as imatinib (Glivec), and compounds that interfere with the function of receptors in the EGFR family, such as trastuzumab (Herceptin), have demonstrated that the paradigm of receptor targeting can be extended beyond steroid-hormone receptors (Pollak et al., 2004). Therapeutic strategies targeting the IGF-1R by eliminating it from the cell membrane, blocking the interaction with IGFs, or interrupting the signal transduction pathway downstream of IGF-1R, may therefore be an effective broad-spectrum anticancer action (Yu and Rohan, 2000; LeRoith et al., 1995).

3.2 Goals of the study

The IGF system is essential for effective embryonic and postnatal growth. Malfunction of the IGF system leads to many serious diseases, including growth and stimulation of cancer cells. In order to manipulate the IGF system in the treatment of diseases connected with IGF deregulation, protein-protein interactions among the elements of the IGF system at the molecular level must be understood. A precise model of the IGF-1R activation by ligands can be created based on high-resolution three-dimensional structures of the interacting proteins or information on their dynamic behaviour from NMR studies.

The main goal of the work presented in this part of the thesis was to determine the molecular basis of biological actions of IGF-1R, especially the interaction with IGF1, using various biochemical techniques, NMR spectroscopy and x-ray crystallography.

3.3 Results

Structural analysis of IGF-1R is important for understanding its precise molecular mode of ligand binding. The main goal of the project was the characterization the IGF-1R structural and biochemical properties, beginning from establishing its effective production and purification. Studies presented in this subsection were carried out on different receptor domains and their combinations. They comprise numerous trials of expression and purification of IGF-1R fragments.

3.3.1 Construct design and cloning

The domain organization of the studied IGF-1R is shown in Figure 3.3.1. All constructs used in this work were designed based on the already known parts of the IGF-1R and IR structures, their comparison (Lou et al., 2006) and alignment the amino acid sequence of IGF-1R to the one of the IR. IR is extensively characterized in the published literature (McKern et al., 2006; De Meyts 2008). Crystal structures of an inactive segment of the first three domains (L1-CR-L2) of the IGF-1R (Garrett et al., 1998) and the ectodomain of the IR without insulin (McKern et al., 2006) have been solved in 1998 and 2006, respectively.

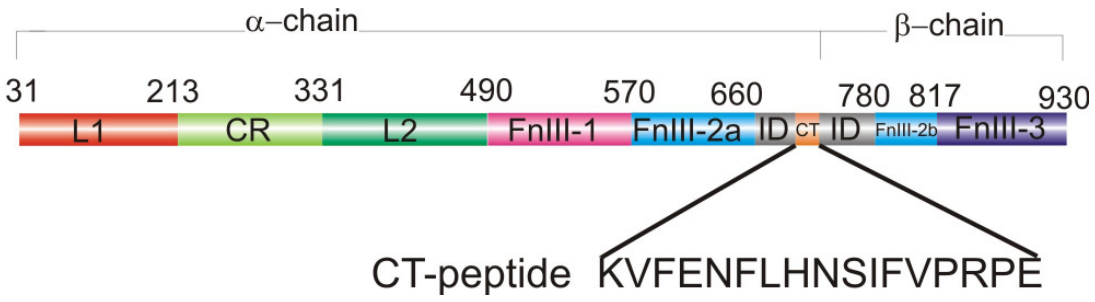

Figure 3.3.1. Schematic representation of the domain organization of IGF-1R. L1 and L2, leucine-rich repeat domains; CR, the cysteine-rich region; FnIII-1, FnIII-2a/2b, FnIII-3, fibronecting type III domains; ID, the insert region of FnIII-2; CT, the so called CT-peptide, the C-terminal 17 residues of the α-chain.

Based on the domain organization of IGF-1R and particular function of each subunit a number of constructs were prepared, expressed, and purified.

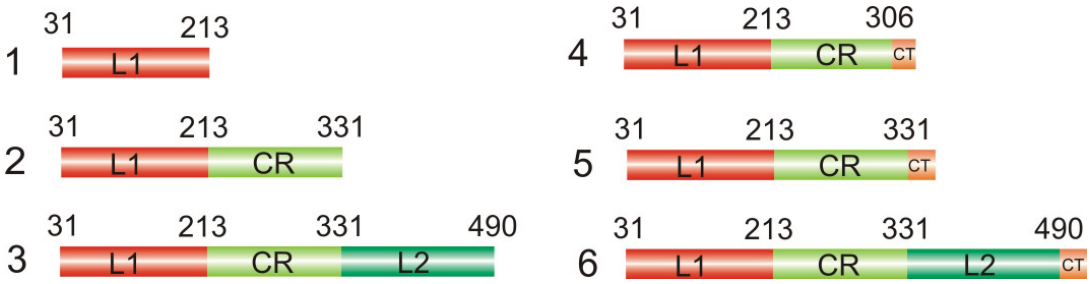

Figure 3.3.2. Constructs of IGF-1R designed for the expression (1-6);
(4-6) – constructs with the CT peptide covalently attached to the C-terminus.

All IGF-1R domains that were subject of this study were cysteine rich and had thus multiple disulfide bonds. This principle feature strongly influenced protein expression and the purification pathway. Therefore two diverse constructs of L1CR domains which contained different number of disulfide bonds were designed.

The so called CT-peptide, being the C-terminal 17 residues of the α-chain was included in some of the constructs due to its association with IGF binding by IGF-1R.

At the beginning constructs were cloned from cDNA purchased from the BD Biosciences Clontech (USA). DNA sequencing revealed that the CR domain of IGF-1R contained a mutation which caused shifting of a reading frame in domains CR and L2. Next, cloning from the pIRESneo2 vector containing full-length IGF-1R received from Roche gave the same result. Finally, Hela cDNA received from collaborators (Jentsch, 2007) was used as a template for construct cloning and optimization.

Because of difficulties first with the expression and later with purification of reasonable amounts of a soluble protein, many different strategies were tried. For this purpose a large number of different vectors were used. For the *E. coli*

expression, constructs were cloned into several vectors pET series resulting in intracellular as well as with extracellular expression of variably tagged proteins.

Table 3.3.1. List of constructs of IGF-1R prepared for the expression

Construct	Destination					
	E.coli		Sf9 (insect)		*Pichia pastoris*	
	Primer	Vector	Primer	Vector	Primer	Vector
L1(31 – 213)	1,2	pET44				
	1,2	pET46				
L1CR(31 – 306)	1,3	pET44		pIEx-7*		
	1,3	pET46				
	7,8	pETM-20				
	7,8	pETM-50				
	7,8	pETM-41				
	7,8	pET20b				
	7,8	pET40b				
L1CR(31 – 331)	1,4	pET44		pIEx-7*	12,17	pICZB
	1,4	pET46			11,13	pICZαC
	7,9	pETM-20				
	7,9	pETM-50				
	7,9	pETM-41				
	7,9	pET20b				
	7,9	pET40b				
L1CRL2(31 – 490)	1,5	pET44		pIEx-7*	12,17	pICZB
	1,5	pET46			11,13	pICZαC
	7,10	pETM-41				
	7,10	pET20b				
	7,10	pET40b				
L1CR(31 – 306)CT	1,6	pET46			12,17	pICZB
					11,13	pICZαC

* constructs received from collaborators (Sitar, 2007)

For yeast expression, constructs were based on vectors from the pICZ series. Constructs in the pIEX vector for the baculovirus expression system were obtained from collaborators (Sitar, 2007). All constructs were cloned using standard protocols described in section 2.2.2. The list of the constructs of IGF-1R is shown in Figure 3.3.2 and Table 3.3.1.

Aggregation was the main problem during purification of the L1CR(31-306) and L1CR(31-331) constructs. To increase the solubility of expressed constructs, different fusion tags listed in Table 3.3.2 were tested. The schematic organization of such tagged constructs is depicted in Figure 3.3.3. Samples purified after overnight cultures usually contained only a fusion tag, which suggested strong degradation. In order to avoid this, constructs for periplasmatic expression were prepared (pET20b, pETM50-Dsba).

Table 3.3.2. List of protein tags used.

Vector	Tag	Cleavage site
pETM-20	*N'Trx-tag*	*TEV*
pETM-41	*N'MBP-tag*	*TEV*
pET44	*N'NusA-tag*	*Enterokinase*
pETM-50	*N'DsbA-tag*	*TEV*

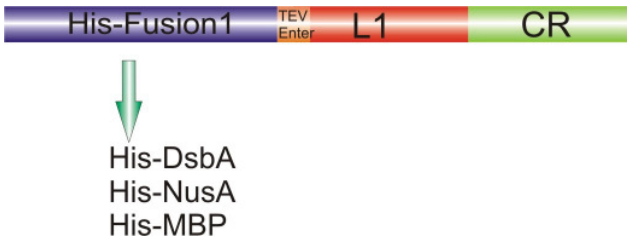

Figure 3.3.3. Schematic representation of the tagged constructs of IGF-1R.

While alternating expression and purification experiments were preformed, it became clear that solubility of the longer constructs (L1CRL2) is higher than of the shorter ones (L1CR). Based on the experience made in our group before and availability of suitable DNA, a construct containing the C-terminal domain of the very soluble cyclase associated protein (CAP) on the C-terminus of L1CR was designed. Structurally CAP is similar to the L2 domain as shown in Figure 3.3.4 (Dodatko at al., 2004). This fusion protein could be later cleaved with Factor Xa.

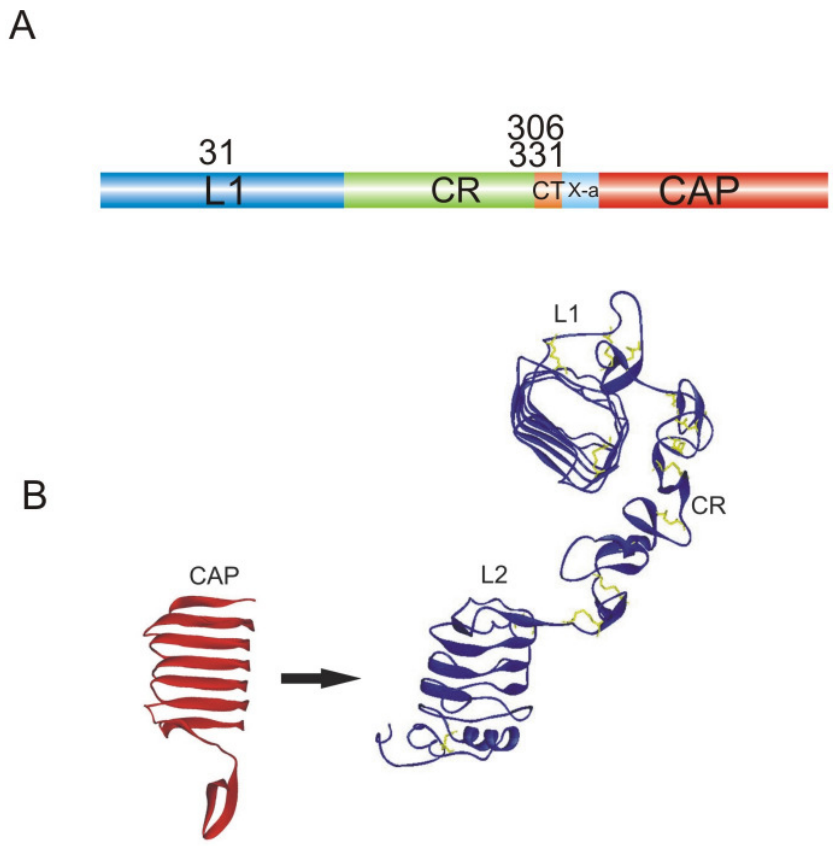

Figure 3.3.4. Construct designed for the expression in *E. coli*. The L2 domain was substituted with CAP in order to increase solubility. CAP is highly soluble and highly expressed in *E. coli.* CAP can be removed from L1CR-CT by cleavage with Factor Xa.

3.3.2 *E. coli* expression and purification

The expression and purification of reasonable amounts of soluble and correctly folded proteins created several problems. In the course of experiments many different strategies were tried. This section contains results of these trials in a logical although not always chronological order. All ideas and tricks employed here came from the examination of already existing methods of purification of similar or homologous proteins.

The information about properties of a protein of interest, such as its molecular mass (MW), number of disulfide bonds, or extinction coefficient (E280) is important in the design of a purification workflow and for other planned investigations. Relevant data were deduced on the basis of the amino acid composition with the aid of ProtParam on www.expasy.ch (Gasteiger et al., 2003). Closer examination of the protein's amino acid sequence and already known structures gave potential cleavage/degradation sites, as well as a secondary structure prediction. Table 3.3.3 contains information on the basic physicochemical properties of the constructs produced or used in the work.

Table 3.3.3. Physicochemical properties of the investigated constructs

	Construct name	N.of aa	MW (Da)	E_{280} ($M^{-1}cm^{-1}$)	N. of S-S bonds
1	L1(31 – 213)	183	21240	34880	4
2	L1CR(31-306)	276	31426	47215	8
3	L1CR(31-331)	301	34082	48955	10
4	L1CRL2(31 – 490)	460	52233	67655	12
5	L1CR(31-306)CT	293	33481	47215	8
6	L1CR(31-331)CT	318	36137	48955	10
7	L1CRL2(31 – 490)CT	477	55391	67655	12

Expression of the protein was carried out in several bacterial strains and under variable conditions. Choice of the suitable strain, as well as optimization of expression time and temperature, is covered in detail in section 2.2.5.

Purification of IGF-1R constructs was performed under native as well as denaturing conditions. As a rule it employed affinity chromatography, followed by gel filtration. Protease cleavage of the tag was also possible. The general strategy for purification is shown in Figure 3.3.5.

Affinity chromatography was the first step of the purification of IGF-1R. Presence of the His-tag in all the expressed proteins enabled the use of nickel slurry. The proteins obtained after this step of purification were usually pure, but in some cases the samples were a mixture of domains, short peptides, degeneration products and aggregates. The final purification was achieved by gel filtration in the proper buffer for further applications of the protein.

Purification under denaturing conditions must be followed by refolding of the protein. First all disulfide bridges must be reduced. Refolding is initiated by reducing the concentration of a denaturant. This can be performed by one- or multi-step dialyses, buffer exchange by gel filtration, applying a gradient of a refolding buffer when the protein of interest is immobilized on the IMAC matrix (solid phase refolding), or by dilution (Tsumoto et al., 2003). For refolding of the IGF-1R constructs a rapid (pulse) dilution technique, as well as one step dialyses, were applied.

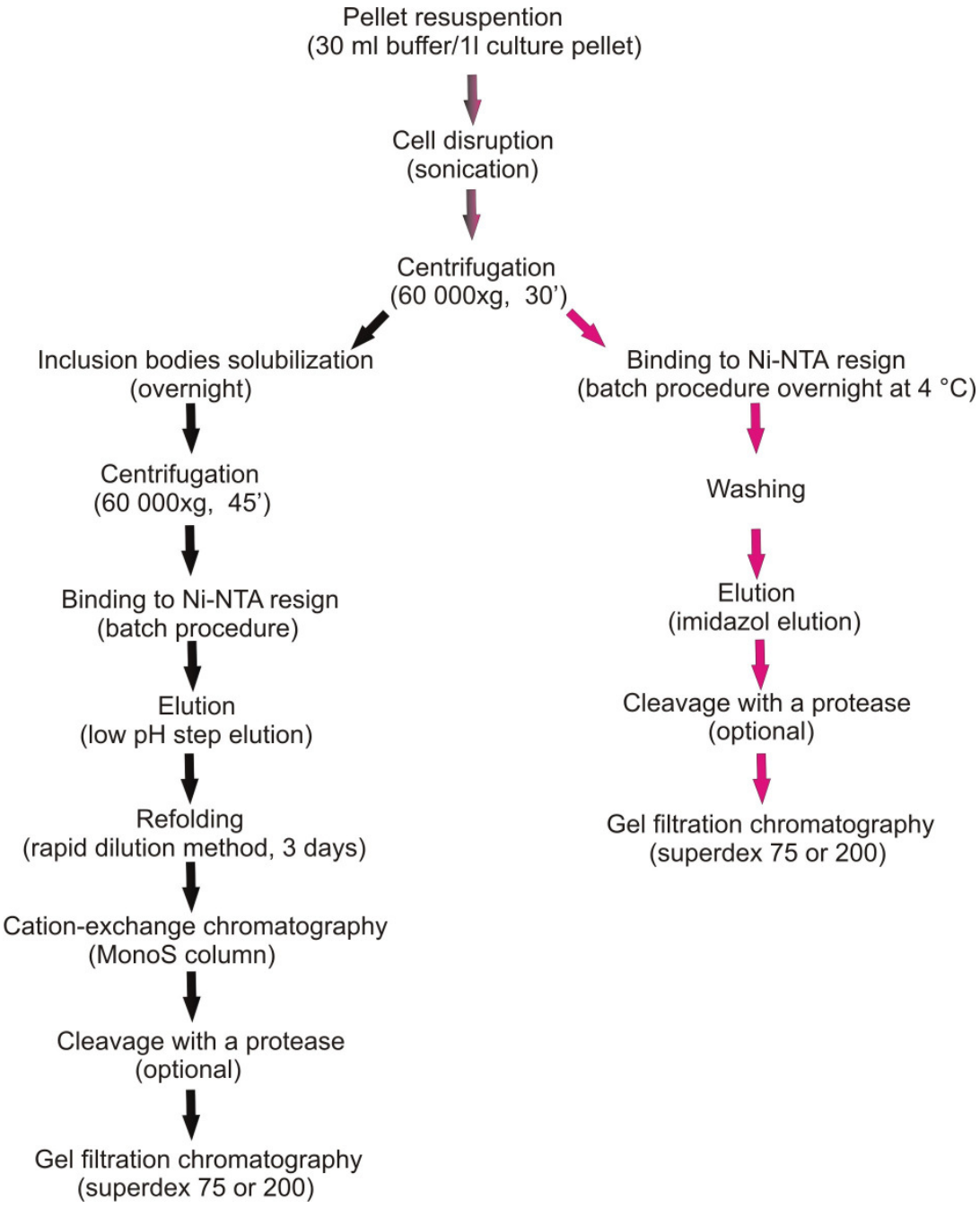

Figure 3.3.5. Flow chart of the purification schemes for IGF-1R constructs. Black arrows follow initial purification under denaturing conditions, red is a native conditions pathway

3.3.2.1 Expression and purification of the L1 domain of IGF-1R

The L1 domain construct was the first one expressed during this study. It was conducted as a pilot expression and purification of prepared constructs. Two vectors and two strains of *E. coli* were tested.

Expression from pET46 in BL21(DE3) CodonPlus gave a large amount of protein, however, only in an insoluble fraction. Initial purification was carried out under native as well as denaturing conditions. Western-blot analysis (His probe) showed that in both cases there was no soluble protein. After solubilization of inclusion bodies in 6M Gu-HCl, the protein was purified on Ni-NTA column. Unfortunately refolding was unsuccessful.

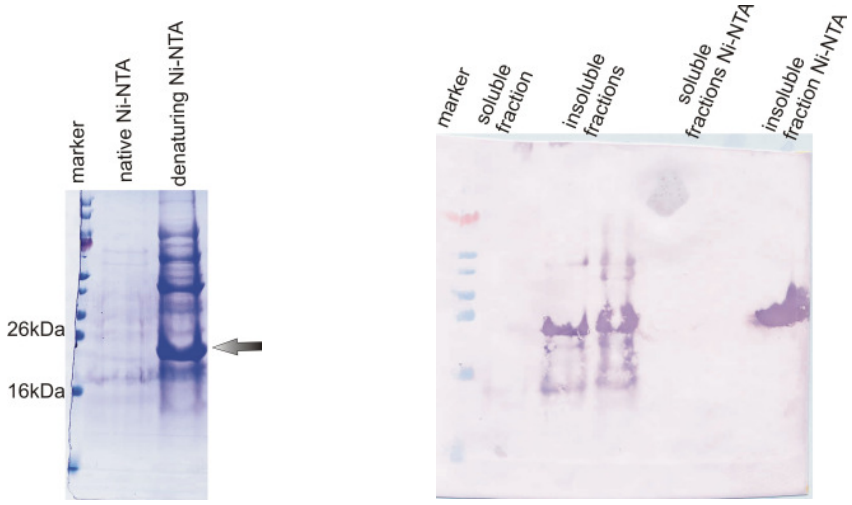

Figure 3.3.6. SDS-Page and Western Blot analysis of the L1 domain expression after binding to the Ni-NTA column. The His- tagged protein was expressed from pET46 in the *E. coli* BL21(DE3) CodonPlus.

Expression from pET44 (His-tag-NusA) in Rosetta-gami (trxB/gor mutant) gave a small amount of the protein in a soluble fraction. The L1 domain, which contains 4 disulfide bonds, was fused to NusA (495-aa), which is a highly soluble protein. pET44 is compatible with the trxB/gor (thioredoxin reductase and glutathione reductase) mutants of *E. coli* strains, which facilitate disulfide bonds

formation in the cytoplasm. Initial purification was performed under native conditions. The yield of the fusion protein obtained was small, but in a soluble fraction.

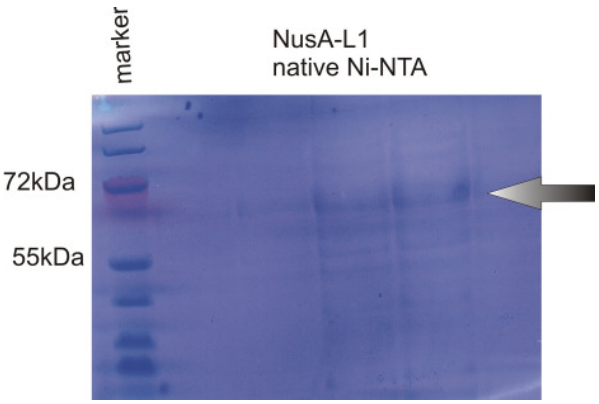

Figure 3.3.7. SDS-PAGE analysis of the L1 domain expression and purification. The His- tagged NusA-fusion protein was expressed from pET44 in *E. coli* Rosetta-gami.

Expression of the L1 domain was performed as check whether any expression of IGF-1R fragments in *E. coli* is possible and for pre-establishing the purification pathway. It gave some positive information, so later the work was focused on the longer constructs, as they are functionally more interesting.

3.3.2.2 Expression and purification of the L1CR(31- 306) domain of IGF-1R

Same as in the case of the expression of the L1 domain, two strains of *E. coli* were used for the initial expression of the L1CR (31 – 306) construct. Expression was performed in pET46 (His-tag) and in both bacterial strains, Rosetta-gami and BL21(DE3), CodonPlus gave a reasonable amount of the protein in a soluble fraction. Worth to mention is that the BL21(DE3) CodonPlus culture resulted in a high expression level and the protein presence in a soluble fraction in two temperatures (15℃ and 37℃), while the Rosetta-gami culture delivered a moderate yield of the protein and a soluble fraction came only from the expression carried out at 15℃.

The protein expressed in Rosetta-gami was purified under non-reducing conditions due to the potential proper disulfide bonds formation in the host cytoplasm during expression. After gel filtration, the protein was pure but aggregated as seen in Figure 3.3.8. Mass spectrometry analysis of the sample collected from peak 1 showed that the sample contained mixture of the 32 kDa protein, which corresponds to the L1CR domain, and short peptides, which were most probably covalently bound to cysteines of L1CR. The molecular mass of the protein in the sample was 41 kDa. Addition of β-mercaptoethanol did not disrupt aggregates.

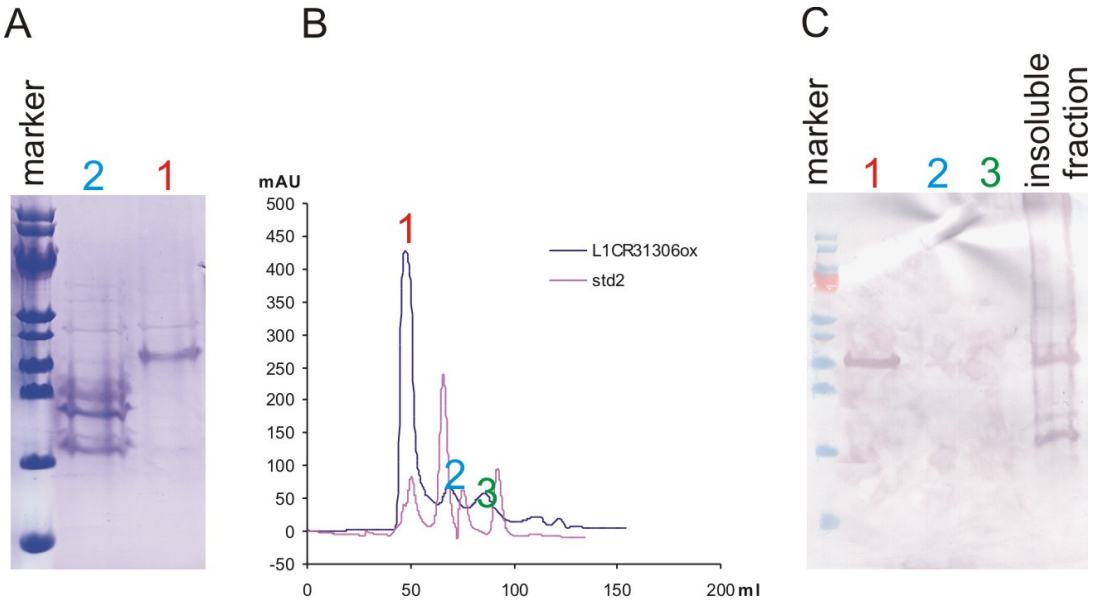

Figure 3.3.8. (A) - SDS-Page analysis of the L1CR(31-306) domain after gel filtration on superdex200. The His-tagged protein was expressed from pET46 in *E. coli* Rosetta-gami (trxC/gor mut);
(B) - elution profile of the L1CR(31-306) from superdex200; gel filtration standards from the left: 440 kDa; 158 kDa; 43 kDa; 13 kDa;
(C) - Western-blot analysis of the L1CR(31-306) domain expression and purification.

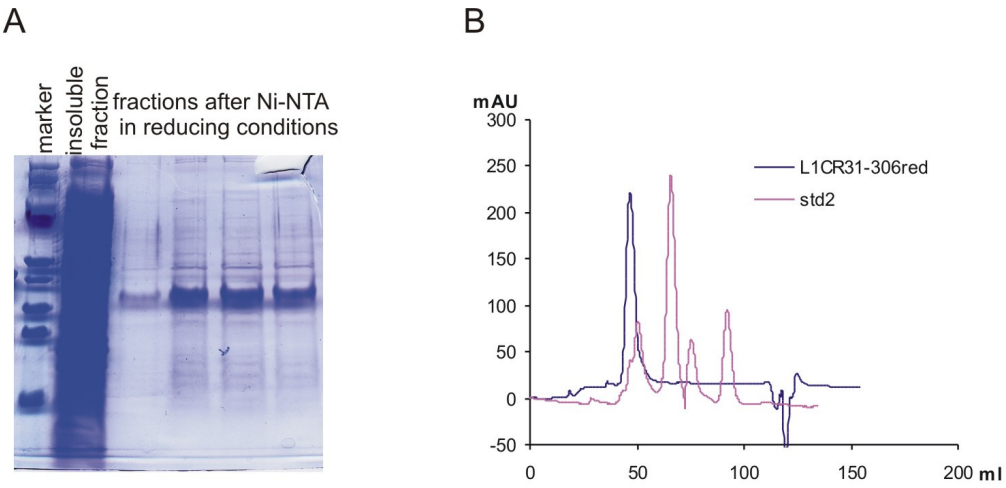

Figure 3.3.9. (A) - SDS-Page analysis of the L1CR(31-306) domain expression after binding to the Ni-NTA column in reducing conditions. The His-tagged protein was expressed from pET46 in *E. coli* BL21(DE3) CodonPlus;
(B) - elution profile of the L1CR(31-306) from superdex200; gel filtration standards: (from the left: 440 kDa; 158 kDa; 43 kDa; 13 kDa).

3.3.2.3 Expression and purification of the L1CR(31- 331) domain of IGF-1R

L1CR(31-331) was expressed from pET44 and pET46 vectors both in BL21(DE3) CodonPlus and Rosetta-gami strains. The expression level in a soluble fraction was higher in BL21(DE3) CodonPlus for both vectors. Expression in Rosetta-gami, especially from pET44 vector, was hardly detectable.

L1CR(31-331) expressed in the BL21(DE3) CodonPlus was purified under variable native and reducing conditions. The composition of buffers is covered in details in section 2.1.12 (Materials and Methods). Unfortunately, the protein expressed from the pET46 vector was in all cases strongly aggregated. The NusA-fusion protein did not aggregate.

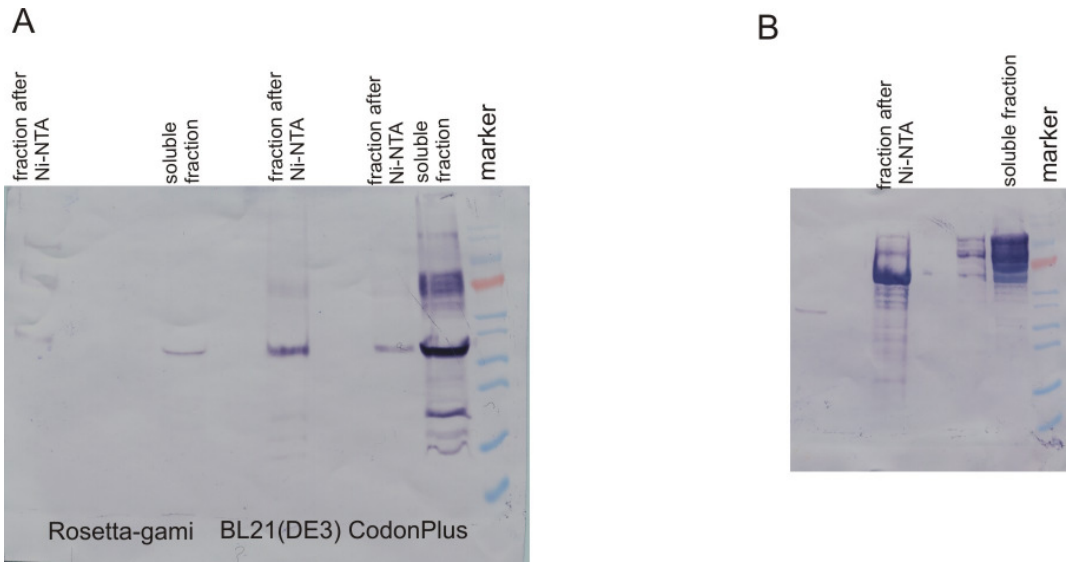

Figure 3.3.10. (A) - Western-blot analysis of the expression of the L1CR(31-331) domains. The His-tagged protein was expressed from pET46 in *E. coli* BL21(DE3), *E. coli* BL21(DE3) CodonPlus and Rosetta-gami and purified under reducing conditions;
(B) - Western-blot analysis of the expression and purification of the L1CR(31-331) domains. The His-tagged NusA-fusion protein was expressed from pET44 in the *E. coli* BL21(DE3) CodonPlus and purified under reducing conditions.

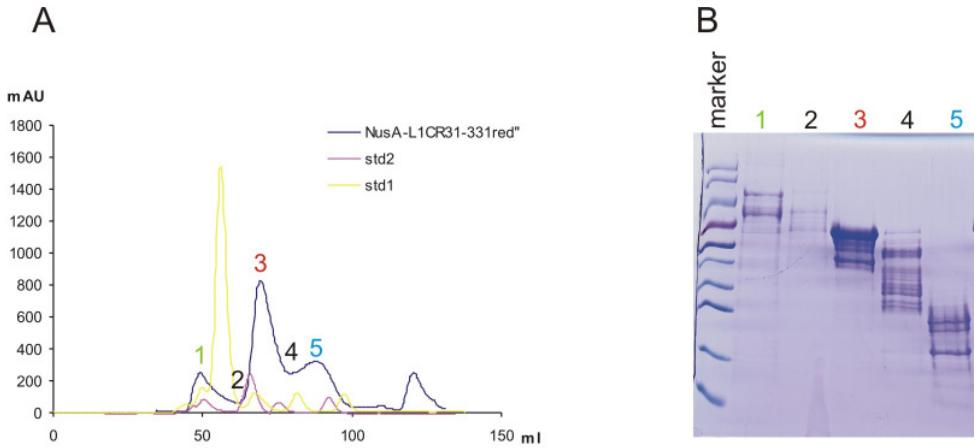

Figure 3.3.11. (A) - Elution profile of the L1CR(31-331) from superdex200; gel filtration standards: (yellow from the left: 440 kDa; 158 kDa; 43 kDa; 13 kDa; violet from the left: 600 kDa; 258 kDa; 67 kDa; 25 kDa);
(B) - SDS-PAGE analysis of purification for the His-tagged NusA-fusion L1CR(31-331) domain.

The NusA-L1CR(31-331) construct was initially purified on gel filtration under reducing conditions. This step was followed by dialysis to non-reducing buffer in order to reoxidize all disulphide bonds.

The reoxidized NusA-L1CR(31-331) was cleaved with enterokinase. The obtained protein was completely degraded.

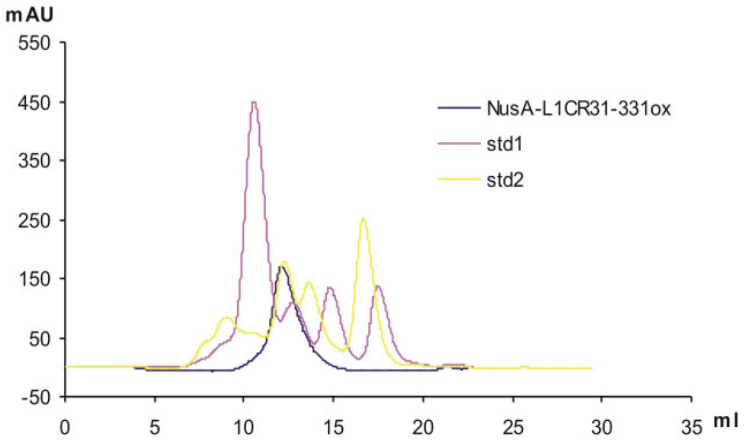

Figure 3.3.12. Elution profile of the L1CR(31-331) from superdex200 after reoxidation; gel filtration standards: (violet from the left: 440 kDa; 158 kDa; 43 kDa; 13 kDa; yellow from the left: 600 kDa; 258 kDa; 67 kDa; 25 kDa)

3.3.2.4 Expression and purification of the L1CRL2(31- 490) domain of IGF-1R

The L1CRL2(31-490) was, identically to L1CR(31-331), expressed from pET44 and pET46 vectors both in the BL21(DE3) CodonPlus and Rosetta-gami strains. The expression level in a soluble fraction was higher in BL21(DE3) CodonPlus for both vectors. Expression in Rosetta-gami, especially from pET44 vector, was hardly detectible. It led to the conclusion that purification under reducing conditions followed by reoxidation in non-reducing buffer is more effective than the expression and purification from the Rosetta-gami strains.

After initial purification, the protein expressed from the pET44 vector was in all cases strongly degraded. The yield of protein for constructs produced in pET46 was acceptable.

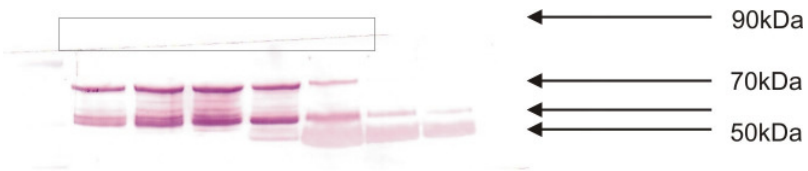

Figure 3.3.13. Western-blot analysis of the expression of the L1CRL2(31-490) domains. The His-tagged protein was expressed from pET44 in *E. coli* BL21(DE3) CodonPlus and purified under reducing conditions. Degraded fragments of the protein with NusA visible as dark bands and indicated by arrows. The full length L1CRL2 would appear in the position marked by the rectangle.

Next, the L1CRL2(31-490) produced in pET46 was purified on gel filtration under reducing conditions. This step showed unfortunately that the protein aggregated strongly. The NMR analysis showed that a small not-aggregated amount of protein obtained after purification was unfolded and unstable.

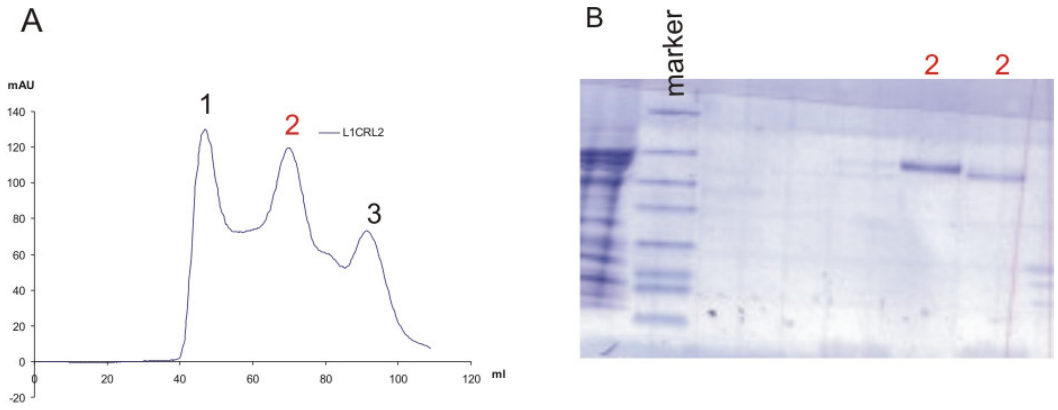

Figure 3.3.14a. (A) - Elution profile of the L1CRL2(31-490) produced in pET46 from superdex200;
(B) - SDS-PAGE analysis of purification for the L1CRL2(31-490) domain.

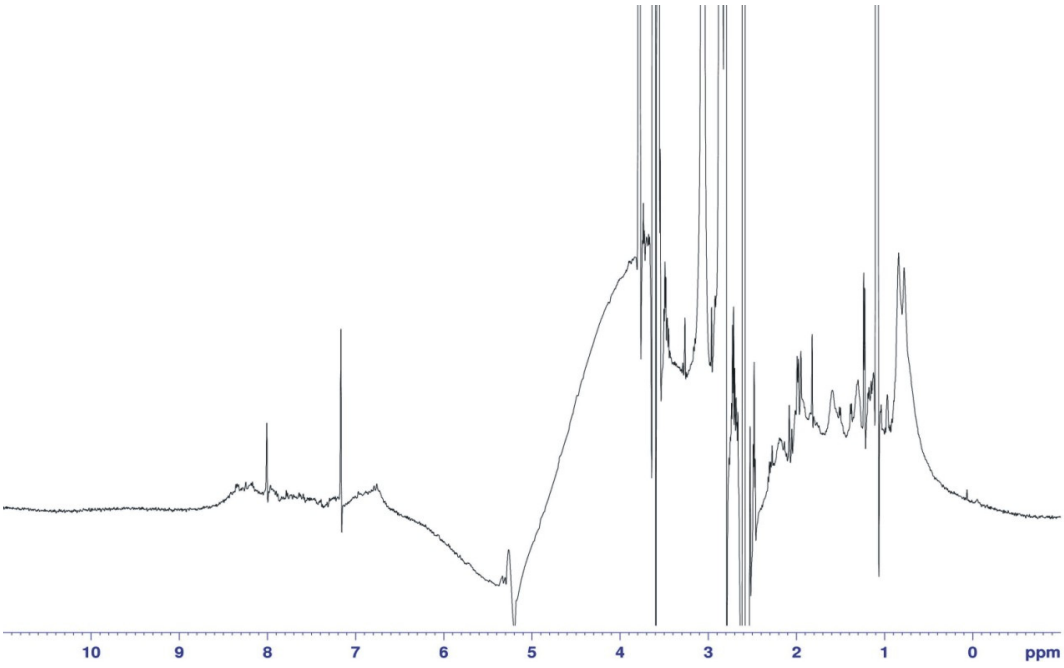

Figure 3.3.14b. 1D NMR spectrum of sample 2 (see Figure 3.3.14a).

3.3.2.5 Purification of the TRX-L1CR(31-306) construct

Aggregation was the main problem during purification of L1CR(31-306) and L1CR(31-331) constructs. As already mentioned, to increase the solubility of the expressed construct, fusion tags were employed. All prepared constructs are listed in Table 3.3.2. Expression of Trx-L1CR(31-306) from the pETM-20 vector was successfully carried out. Samples obtained from pETM-41 (MBP- tag) and pETM-50 (DsbA-tag) vectors, purified after overnight cultures, usually contained only the fusion tag, which suggested strong degradation of the IGF-1R proteins.

Unfortunately, also in this case the protein obtained was unstable and degraded.

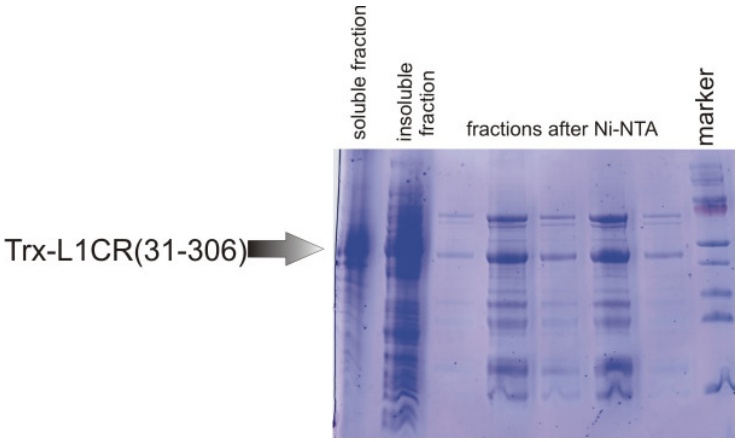

Figure 3.3.15. SDS-PAGE analysis of the expression and purification in reducing conditions of the His-Trx L1CR(31-306) domain. His-tagged protein was expressed from pETM-20 in the *E. coli* BL21(DE3) CodonPlus;

3.3.2.6 Purification of the L1CR306CTCAP construct

During the so far conducted experiments it was noticed that the solubility of the longer constructs (L1CRL2) is higher than that of shorter ones (L1CR) so the construct containing a very soluble protein (CAP), which is similar to the L2 domain, was designed at the C-terminus of L1CR. The fusion protein could be cleaved with Factor Xa. Protein was expressed in pET44 (only NusA fusion protein was purified) and pET46 in both BL21(DE3)Codon Plus and Rosetta-gami. The protein expressed from pET46 in BL21(DE3)Codon Plus was purified in reducing conditions and then reoxidized. Expression, purification and the NMR spectrum of the L1CR306CTCAP are shown in Figures 3.3.16 and 3.3.17. The protein was still aggregated. Additionally, after cleaving with Factor Xa, the protein was completely degraded.

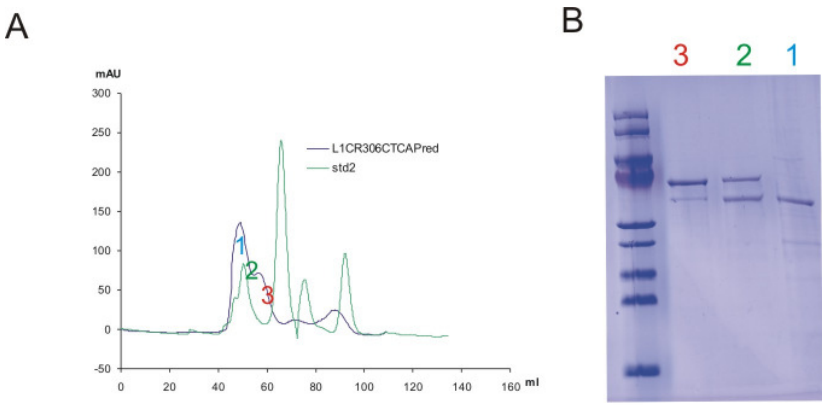

Figure 3.3.16. (A) - elution profile of the His-L1CR(31-306)CTCAP from superdex200 in reducing conditions; gel filtration standards: (from the left: 660 kDa; 258 kDa; 67 kDa; 25 kDa);
(B) - SDS-Page analysis of the His-L1CR(31-306)CTCAP purification.

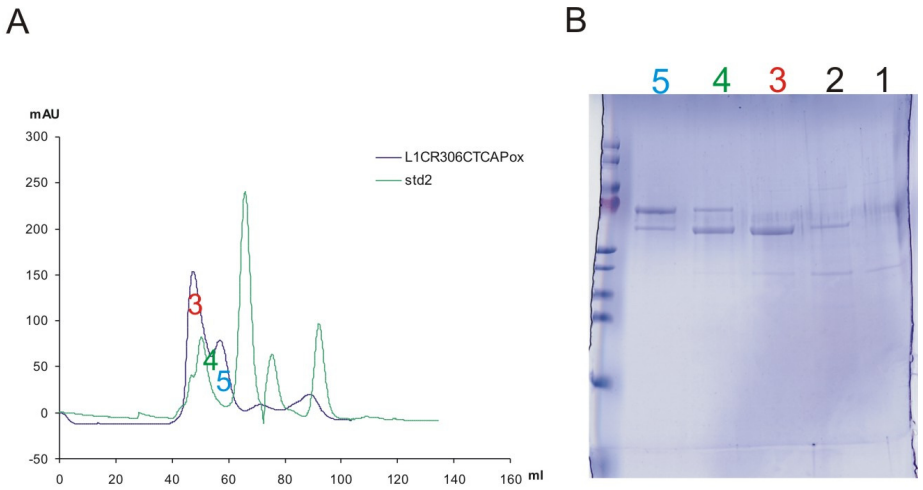

Figure 3.3.17a. (A) - elution profile of the His-L1CR(31-306)CTCAP from superdex200 after reoxidation by dialysis to a non-reducing buffer; gel filtration standards: (from the left: 660 kDa; 258 kDa; 67 kDa; 25 kDa);
(B) - SDS-Page analysis of the His-L1CR(31-306)CTCAP purification

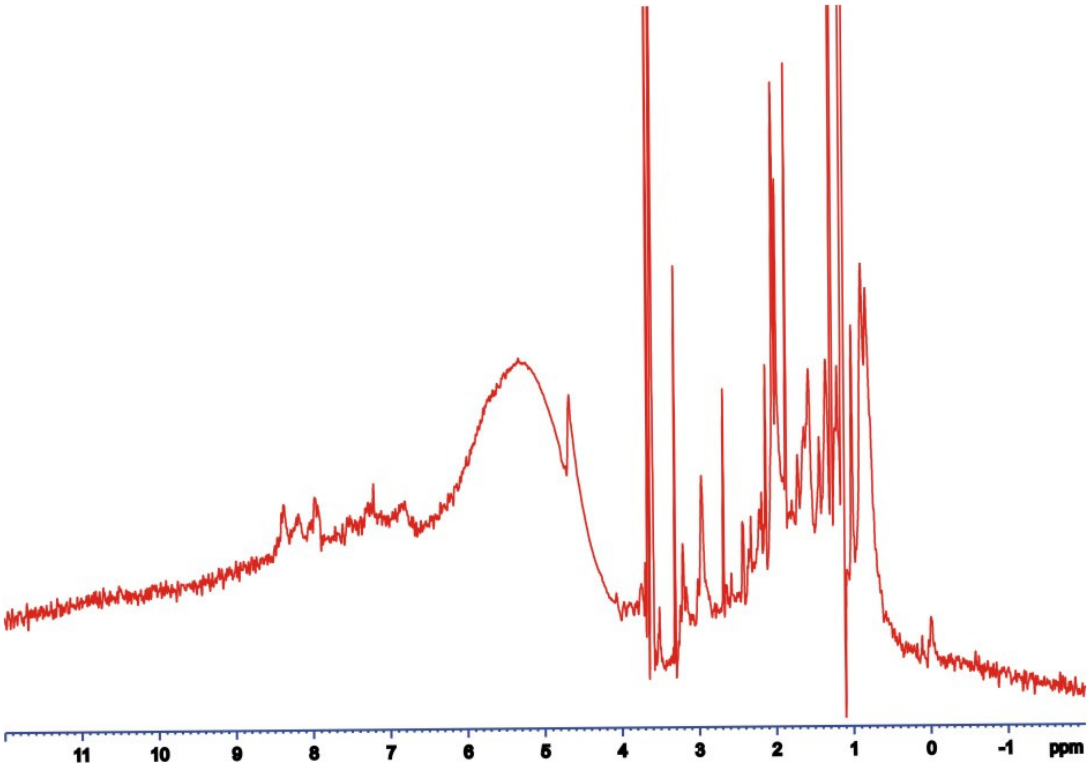

Figure 3.3.17b. 1D NMR spectrum of sample 3.

3.3.2.7 Periplasmatic expression of the L1CR(31-306) domain of IGF-1R

In order to avoid aggregation and degradation of expressed IGF-1R domains, the periplasmatic expression was attempted. Constructs of the L1CR(31 – 306) domain in pET20b and pET40b were expressed to the periplasm in the BL21(DE3) Codon Plus; BL21(DE3) - pGKJE8 and Tuner bacterial strains.

In case of the construct in pET20b, the expression level was low and the obtained protein was partially degraded. Expression from pET40b, as a fusion protein with DsbC and N- and C- terminal His-tagged, gave a bit higher yields but SDS-PAGE analysis after purification on Ni-NTA showed two main bands recognised by anty-His antibodies, which suggested the degradation in the middle of the IGF-1R domain.

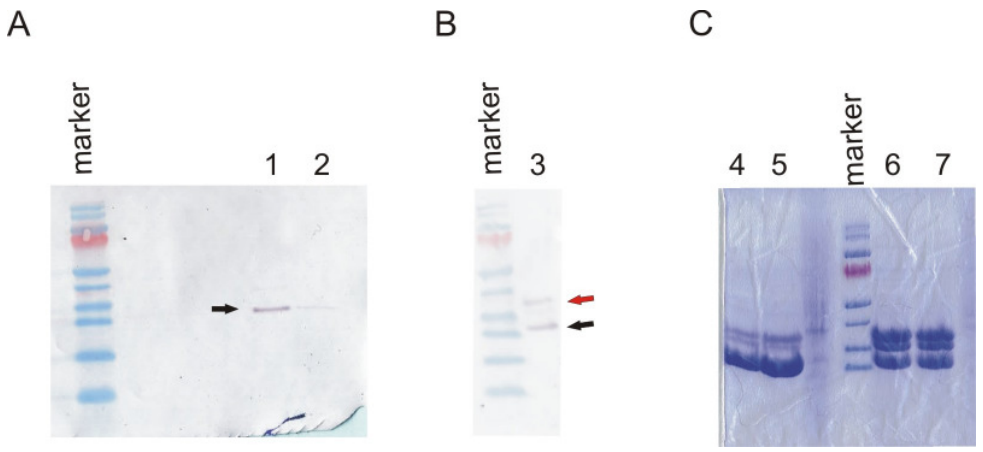

Figure 3.3.18. An example of the Western-blot analysis of purified periplasmic fractions (1, 2, 3).
(A) - L1CR(31-306) indicated by a black arrow was expressed in the *E. coli* Tuner strain from pET20b (C-terminal His-tag).
(B) - pET40b (N- and C-terminal His-tags) SDS-PAGE analysis after purification on Ni-NTA showed two main bands (red and black arrows), recognised by anti-His antibodies which suggests the degradation in the middle of the IGF-1R.
(C) - SDS-PAGE analysis of periplasmic fractions of L1CR(31-306) from pET40b expressed in BL21(DE3) Codon Plus (4, 5) and BL21(DE3) - pGKJE8 (6, 7).

3.3.2.8 Refolding attempts of the L1CR(31-306)CT construct of IGF-1R

The L1CR(31-306)CT construct of IGF-1R in the pET46 vector was expressed in the ArcticExpress RIL strain in 37℃ in order to increase the expression level. The protein was expressed as insoluble inclusion bodies in a 4 l bacterial culture and purified on the Ni-NTA slurry under denaturing conditions. Partial degradation, which was observed in the first trial, was later significantly decreased by shorter time of culture growth after induction (2 h). The amount of the protein purified under denaturing conditions was sufficient for refolding trials.

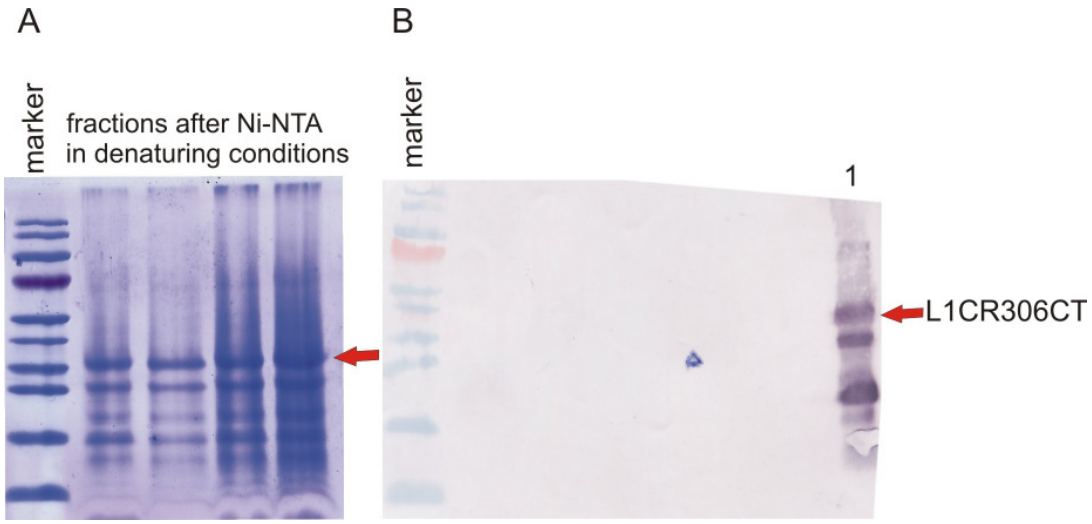

Figure 3.3.19. (A) - SDS-PAGE analysis of the first trial of the expression of the L1CR306CT construct in the ArcticExpress RIL pRARE. Partial degradation is visible.
(B) - Western blot analysis of the purified protein. The expressed IGF-1R construct is indicated by red arrows.

Several refolding buffers and different conditions were tested based on the protocols from the Monash University (http://refold.med.monash.edu.au). Refolding of the purified protein was carried out by dilution, as well as dialysis against refolding buffer. Refolding by dilution was more effective and gave a small amount of a soluble protein with the proper size. Results of refolding by dilution are summarized in Table 3.3.4.

Refolded protein was loaded on an ion exchange column MonoQ. Then, the samples were analysed by mass spectrometry and NMR.

Table 3.3.4. Refolding of the L1CR(31-306)CT construct expressed in the Arctic Express RIL

Refolding strategy	Result
Dilution to buffer containing: 0,2 M L-arginine, 1 mM EDTA, 100 mM Tris, 2 mM GSH, 2 mM GSSG, 10% glycerol	low refolding efficiency; gel filtration suggests aggregation
Dilution to buffer containing: 0,2 M L-arginine, 1 mM EDTA, 100 mM Tris, 5 mM GSH, 1 mM GSSG, 10% glycerol	small amount of soluble protein; NMR analysis suggests partial folding, mass spectrometry has shown aggregation caused by intermolecular disulphide bonds
Dilution to buffer containing: 0,2 M L-arginine, 1 mM EDTA, 100 mM Tris, 10 mM cysteamine, 5 mM cystamine, 10% glycerol	no soluble protein
Dilution to buffer containing: 0,5 M L-arginine, 1 mM EDTA, 100 mM Tris, 5 mM GSH, 1 mM GSSG, 10% glycerol	higher refolding efficiency than in previous trials, sample precipitated during concentration for further analysis

.

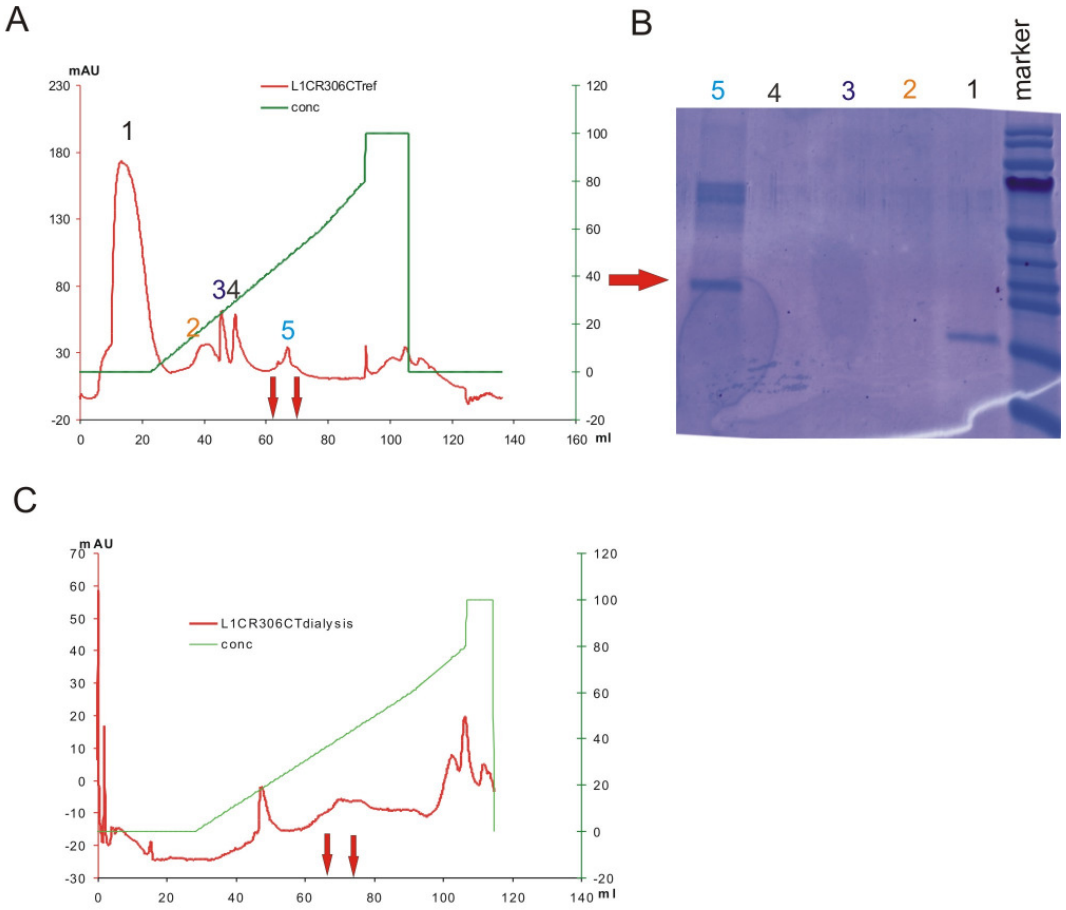

Figure 3.3.20. (A) – Elution profile of L1CR(31-306)CT after the MonoQ column. The protein was refolded by dilution;
(B) - SDS-PAGE analysis of the purification of L1CR(31-306)CT
(C) - Elution profile of L1CR(31-306)CT after a MonoQ column. The protein was refolded by dialysis. Red arrows indicate L1CR(31-306)CT.

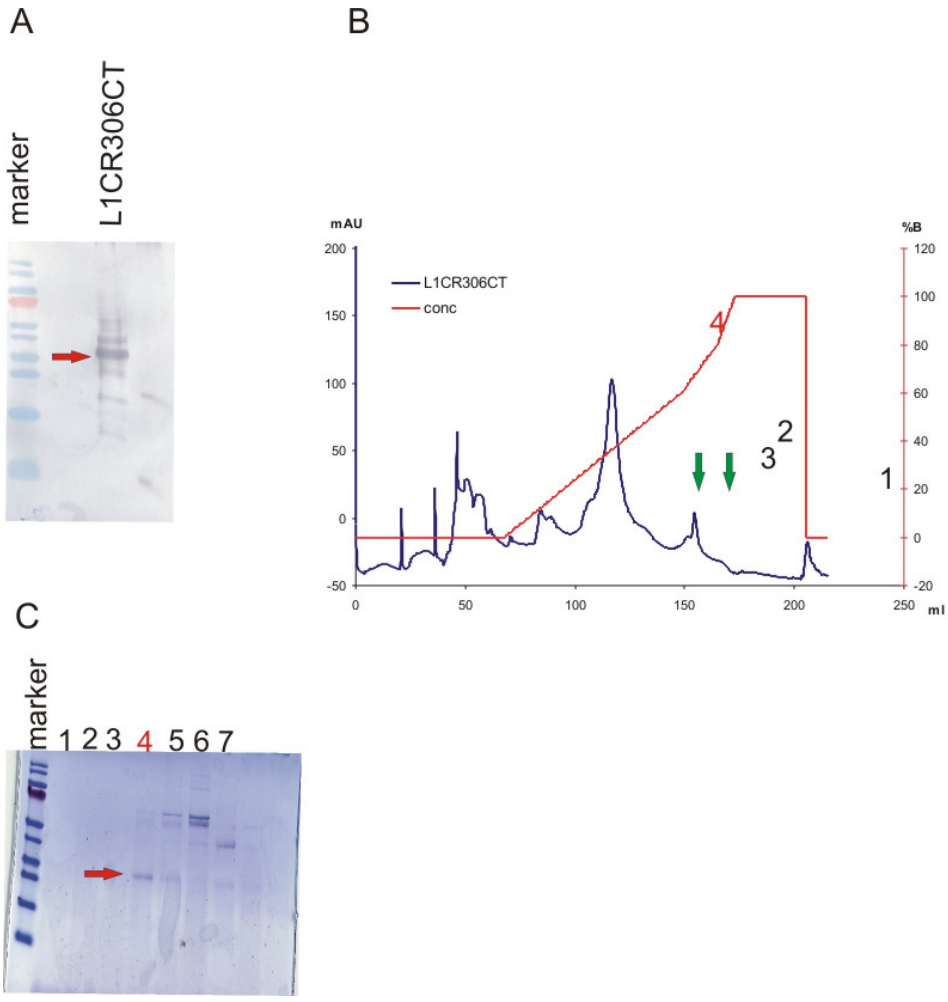

Figure 3.3.21 (A) - Western blot analysis of the refolding of L1CR(31-306)CT expressed in ArcticExpress RIL;
(B) - Elution profile of refolded L1CR(31-306)CT after a MonoQ column. Pooled fractions are indicated by green arrows;
(C) - SDS-PAGE analysis of the purification of L1CR(31-306)CT. Red arrow indicates the construct.

3.3.2.9 NMR functional analyses of the refolded L1CR(31-306)CT construct

All refolded L1CR(31-306)CT constructs were examined by NMR in order to assess their structural integrity necessary for crystallization studies. The unique strength of NMR lies in its capability to estimate unstructured regions of the polypeptide chains in the partially folded proteins and to identify samples that are heterogeneous because of aggregation or other conformational effects

(Rehm et al., 2002). The signal dispersion and line shape of the resonances is dependent on the folding. The line widths of the signal in any NMR spectrum are correlated to the size of the molecule. The NMR signal of large biomolecules decays much faster than that of smaller ones, which produces broader NMR signals. Figure 3.3.22 shows typical one-dimensional proton NMR spectra of the refolded L1CR(31-306)CT constructs.

The constructs of IGF-1R yielded poor NMR spectra, most probably because of the large size and dilution of the protein. Regardless of refolding conditions, samples gave, as a rule, large and broad signals at approximately 8.3 ppm and showed a small dispersion of the amide backbone chemical shift. This suggested unfolding and aggregation. Spectra of the L1CR(31-306)CT construct in oxidizing conditions were 'better' but still indicated that this protein was only partially folded and contained large unstructured fragments. Despite the poor features of NMR spectra of IGF-1R constructs, further structural studies, including crystallization trials, were undertaken.

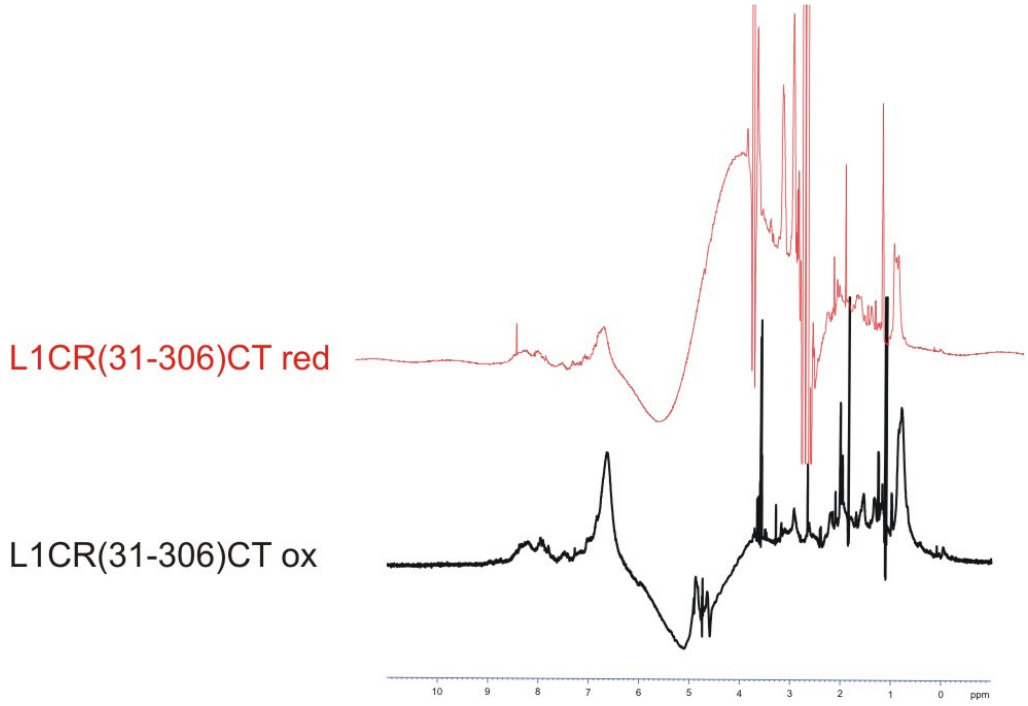

Figure 3.3.22. Characterization of the refolded L1CR(31-306)CT construct of IGF-1R by one-dimensional proton NMR spectra.

3.3.3 Insect cells expression

An expression test of L1CR(31-306) was carried out by collaborators in Sf9 insect cells (Sitar, 2007). There was no protein detectable by Western-blot.

3.3.4 Yeast expression and purification

Two strains of *Pichia pastoris* have been used for expression of the L1CR(31-306)CT and L1CRL2(31-490) constructs KM71H and protease deficient SMD1168H. In the first experiment small amounts of the protein expressed from pICZαC were detected in media by the SDS-PAGE and Western blot. Unfortunately, scaling up of the expression did not give any protein. Later, intracellular expression of IGF-1R constructs from pICZB was performed. Results of this expression were hardly detectable by Western blot and purification afterward from a soluble fraction did not give any protein.

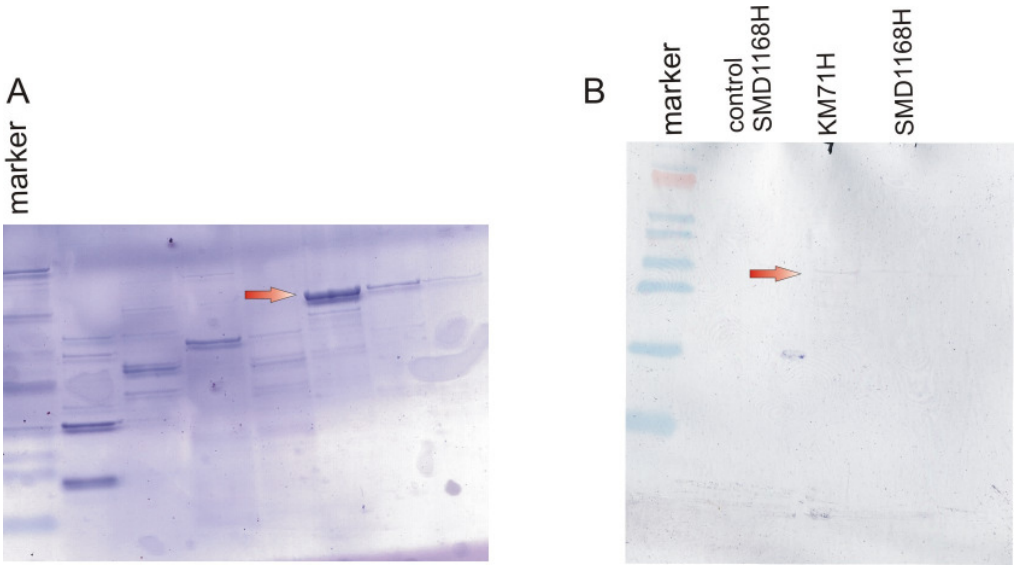

Figure 3.3.23. (A) – SDS-PAGE analysis of the purification of the L1CRL2(31-490) construct expressed extracellular from pICZαC. Media was collected and loaded on the ion exchange column MonoQ.

(B) - Western blot analysis of the intracellular expression of the L1CR(31-306)CT construct in *Pichia pastoris* SMD1168H and KM71H. Small bands with molecular mass corresponding to IGF-1R construct are indicated by red arrow.

3.4 Discussion

IGF-1R and IR show high sequence homology and their structures are similar although the physiological functions of insulin and IGFs are very different. They are both large transmembrane glycoproteins, consisting of two extracellular α-subunits which contain the ligand-binding site and two β-subunits which contain the hydrophobic transmembrane domains (LeRoith et al., 1995; Yamasaki et al., 1993; McKern et al., 2006). The structure of the first three domains of the extracellular portion of IGF-1R has been already determined (Garrett et al., 1998), as well as the crystal structure of the IR ectodomain which has been resolved recently (McKern et al., 2006). The first there domains of the IGF-1R differ structurally from the IR in the regions governing ligand specificity (Lou et al., 2006). However both solved structures didn't show the bound ligand, so the crystallized segments were in their inactive form.

The studies presented in this subsection of the thesis had as a long term goal to find the three-dimensional structure of the extracellular part of IGF-1R in an active form with the bound ligand. Based on the already known IGF-1R and IR structures, their comparison (Lou et al., 2006) and alignment of the amino acid sequence of IGF-1R to the one of the IR, as well as employing the experience made with expression, purification and crystallization of other recombinant proteins, numerous trials of expression and purification of IGF-1R fragments were carried out to obtain it for crystallization. All the constructs used in this work are presented in Figures 3.3.1 and 3.3.2, and in Table 3.3.1 in the previous chapter. The general organization of the receptor domains is shown in Figure 3.4.1 illustrated by the known IR structure.

The main problem of the purification was high aggregation of the protein. In an attempt to increase the solubility of the expressed constructs, expressions with different fusion tags were tested. Also while alternating expression and purification experiments were preformed, it became clear that the solubility of the longer constructs (L1CRL2) is higher than that of the shorter ones (L1CR). Basically, none of these trials worked due to either very strong degradation, or

instability of the obtained proteins. Stable constructs on the other hand tended to aggregate. Also expression to the periplasmic did not solve these problems.

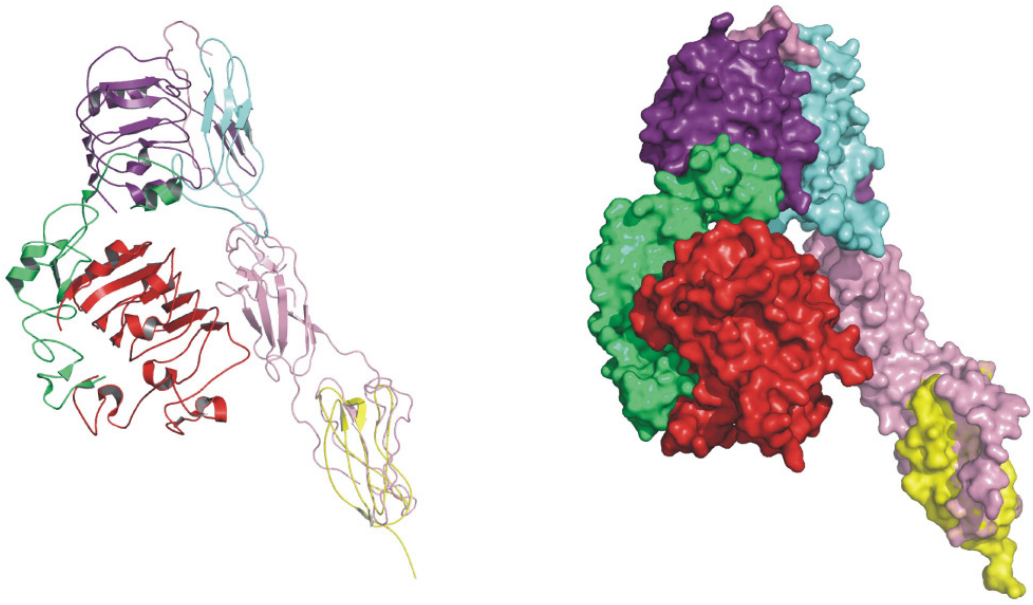

Figure 3.4.1. Overall organization and fold of the IR ectodomain. Left in ribbon, right in surface representation. The L1 domain is in red; CR in green; L2 in violet; FnIII-1 in cyan; FnIII-2 in pink; FnIII-3 in yellow. The L1, CR and L2 domains are from one monomer in the IR homodimer, whereas the FnIII domains are from the other (adapted from pdb file 2DTG, McKern et al., 2006).

All IGF-1R constructs that were subject of this study were cysteine rich and had thus multiple disulphide bonds. The CR region is composed of eight disulphide-bonded modules, seven of which form a rod-shaped domain. Disulphide bonds connect both α- and β-subunits to form a functional heterotetrameric receptor complex (Garrett et al., 1998; McKern et al., 2006). This feature strongly influences protein expression and the purification. A high number of disulphide bonds make it almost impossible to obtain properly folded protein in bacterial expression systems. Correct folding requires a lot of chaperons and protein disulphide isomerases. Otherwise expressed protein aggregates strongly and form inclusion bodies. Expression in insoluble form

followed by refolding was in this case also in vain and did not work. The high number of disulphide bonds decreases the chances of correct refolding of the protein. This disulphide bond problem could be solved by exchanging all cysteine residues by serine at least in the N-terminal part of the CR domain which is not important for binding. Also expression in a system other than the bacterial expression system could increase the probability of arriving at a properly folded protein. Until now trials with yeast or insect cell lines brought minimal yields of protein when compared to test expressions in *E. coli*. Only by means of Western blots the overexpressed protein could be detected. Further advances could be also made by expressing the CR domain in mammalian cell lines. This is a very expensive and tedious undertaking which could not be realized in this study. If this will be tried in the future it has to be taken into account that mammalian expression will yield highly glycosylated protein which does not allow future crystallization. So the chosen cell line must be not only stable but also posses the blocked glycosylation pathway like for example the so called Lec8 mutant CHO cells used by McKern, et al. (1997).

The analysis of the known structures shows the L domains each adopt a compact shape consisting of a single-stranded right-handed β-helix (Garret et al., 1998). The L1 domain possesses a highly hydrophobic exposed region and the interaction between it and the FnIII-2 domain from the second monomer of the receptor has been reported (McKern et al., 2006). These factors must also be taken into consideration. Due to an enormously large number of possibilities of interactions which implicates potential aggregation of the expressed protein, particular domains should be expressed separately. On the other hand, because of the interactions between the domains in the native receptor separate expression of correctly folded domains might be impossible. Only some domains like CR and possibly also FnIII seem to be possible to be expressed and purified separately. According to the known IR structure the L1 domain is surrounded by the FnIII-3 domain of the second monomer of the receptor from the one side and by CR from the other. They cover the highly hydrophobic region exposed in the L1 domain. Hence, further attempts of expression should focus more on

coexpression of these three domains. It would be interesting to test also expression of L2 domain together with the FnIII-1 domian as one construct. These domains are exposed in the structure, so the possibility that they are soluble is higher.

Studies currently under way in our group are directed at expressing the whole ectodomain of the IGF1-R. It is highly possible that some domains need to be in the “native” environment provided by the other domains to enable the correct folding.

4 Mad2

4.1 Biological background

The division of the cell's nucleus (mitosis) is the culmination of the cell cycle. The sister chromatids produced in the earlier phases of the cycle must be segregated to opposite poles of the cell before its division. The aim is to ensure that daughter cells inherit the exact same set of chromosomes possessed by their parent cell and this is accomplished with an accuracy of one error in less than 10^5 events.

The process of chromosome segregation during mitosis is known for nearly 80 years. Although its mechanism seems to be so simple and logical the molecular basis for it is one of the most complicated found in living cells. Each step is subject to precise control where many of the details are still unknown.

One of the crucial regulation mechanisms is the so called mitotic checkpoint (spindle assembly checkpoint) which is responsible for the timing of the beginning of the anaphase in metazoan mitosis. Among several important factors the mitotic arrest deficient 2 (Mad2) protein seems to be the most downstream effector in this process. It is also a possible regulator of the centromeric protein shugoshin (Sgo) (De Antoni et al., 2005; Nasmyth, 2005). Two recently resolved crystal structures of a Mad2 dimer (Yang at al., 2008; Mapelli at al., 2007) have provided new insights into the molecular details of the Mad2 regulation in the spindle checkpoint but a detailed mechanism of the Sgo binding remains unclear.

4.1.1 The mitosis

Mitosis per definition is the process of division of the nucleus of a eukaryotic cell involving condensation of the DNA into visible chromosomes and separation of the duplicated chromosomes to form two identical sets. Mitosis is traditionally divided into five stages (Figure 4.1.1). In the prophase the replicated chromosomes, each consisting of two closely associated sister chromatids,

condense to be subsequently attached to the spindle microtubules during the prometaphase when the nuclear envelope breaks. After the chromosomes are aligned at the equator of the spindle, midway between the spindle poles at the metaphase, during the anaphase sister chromatids synchronously separate from two daughter chromosomes and each is pulled slowly toward the spindle pole. During telophase, the two sets of daughter chromosomes arrive at the poles of the spindle and decondense and a new nuclear envelope reassembles around each set completing the formation of two nuclei and marking the end of mitosis (for a comprehensive overview see Alberts et al., 2002).

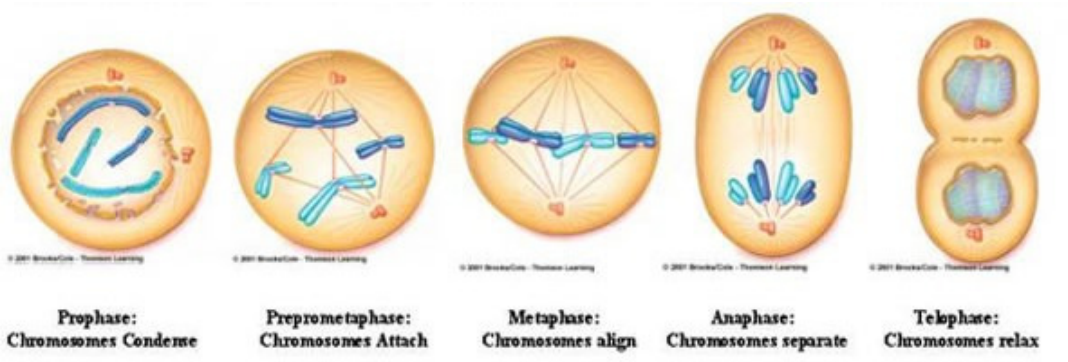

Figure 4.1.1. A schematic overview of the successive stages of the mitosis. Figure adapted from www.biology.iupui.edu.

4.1.2 The mechanism of sister chromatids separation

To prevent the syntelic attachment (attachment of sister chromatids to microtubules emanating from the same pole) and as a result production of aneuploid daughter cells, sister chromatids of replicated chromosomes stay paired until metaphase. This phenomenon, known as sister chromatid cohesion (Miyazaki and Orr-Weaver, 1994) is mediated by a ring-shaped, multisubunit complex, cohesin, which encloses the two DNA double strands in its middle (Gruber et al., 2003; Haering and Nasmyth, 2003; Ivanov and Nasmyth, 2005). The cohesin complex is composed of two Smc-subunits (Structural Maintenance of Chromosomes), Smc1 und Smc3, and at least two other proteins, Scc1/Rad21 and Scc3 (Nasmyth and Haering, 2005). During vertebrate mitosis the vast majority of cohesin is removed from chromosomes in pro- and prometaphase.

This happens in the so-called prophase pathway and requires activity of polo like kinase 1 (Plk1) and aurora B as well as phosphorylation of the Scc3 subunit (Hauf et al., 2005; Sumara et al., 2002). A small centromeric fraction of cohesin stays unresponsive to this process and is not removed until all chromosomes have attached amphitelically and come under tension. This subpolulation of cohesin is protected from the prophase pathway by the centomeric protein called shugoshin, Japanese for a “guardian spirit” (McGuinnes et al., 2005). The remaining cohesin is finally removed in the anaphase by the action of a site specific protease, separase (Figure 4.1.2). This cleaves the Scc1/Rad21 subunit which results in ultimate separation of sister chromatids (Weizenegger et al., 2000; Uhlmann et al., 2000).

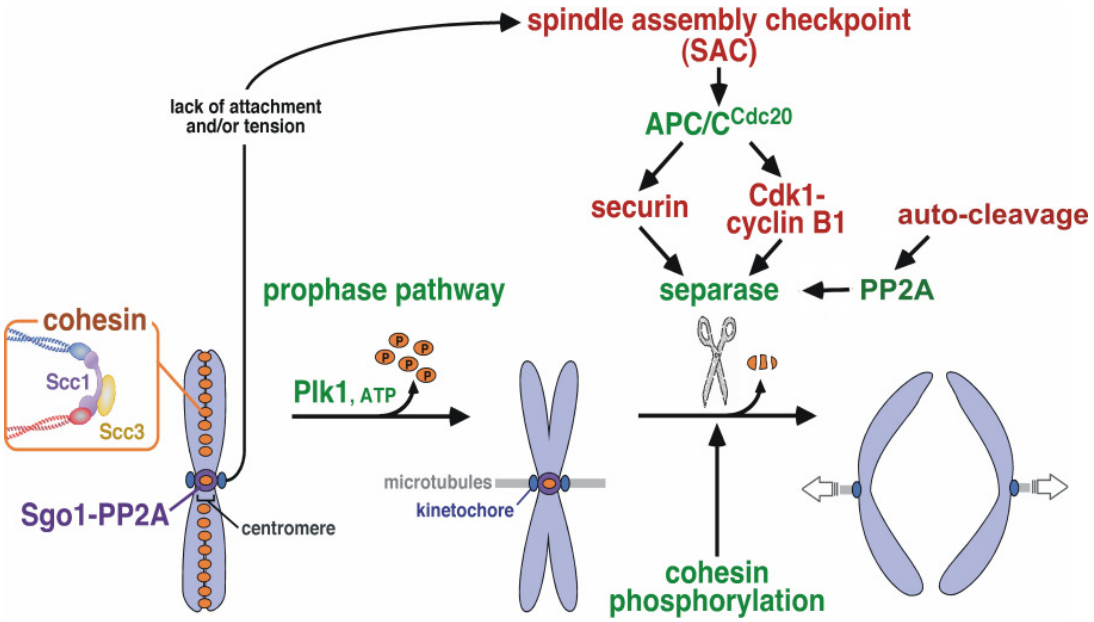

Figure 4.1.2. Current model of sister chromatid separation in vertebrate mitosis. Figure reproduced from Stemmann, (2008). Details regarding the different phases are described in the text.

To avoid that sister chromatids separate before every single chromosome aligns at the equator of the spindle and comes under tension separase is kept inactive until metaphase-to-anaphase transition. In the earlier phases of the cell

cycle separase is associated with an inhibitory chaperon, securin and additionally inhibited due to its phosphorylation by Cdk1/cyclin B1 protein kinase (Nasmyth, 2005). Securin is degraded in late metaphase through the so called ubiquitin-proteasome pathway (Yamamoto et al., 1996). Responsible for this is the anaphase-promoting complex or cyclosome (APC/C), a huge multisubunit ubiquitin protein ligase in conjunction with its accessory protein Cdc20 (Zachariae and Nasmyth, 1999). At the same time APC/C causes destruction of cyclin B and due to this the inhibition of Cdk1 on separase is abrogated and separase becomes active (Boos, 2008; Stemmann, 2008). Active separase cleaves not only the cohesin but also itself, which in turn is regulated by protein phosphatase 2A (PP2A). B' domain of PP2A binds to separase and this way inhibits its autocleavage activity (Kitajima et al., 2006; Holland et al., 2007).

4.1.3 The mitotic checkpoint; Mad2 regulation

The mitotic checkpoint called also spindle assembly checkpoint (SAC) is of prime importance for the timing of anaphase onset in metazoan mitosis. This surveillance mechanism operates to ensure that all chromosomes have aligned on the equatorial plate of the spindle before sister chromatid separation occurs. The checkpoint depends on a sensor mechanism that monitors the state of kinetochore and as long as not all chromosomes are properly attached to the spindle generates a negative signal, blocking the activation of APC/C^{Cdc20} (Rieder et al., 1994).

Among several other proteins (e.g. Bub1 and Bub3) which are recruited to unattached kinetochores and are required for the mitotic checkpoint to function Mad2 seems to be the most downstream effector (Shah and Cleveland, 2000; Musacchio and Salmon, 2007). It exists in two conformations called open (o) or closed (c) form (Figure 4.1.3). In its active (c) conformation Mad2 is capable of binding Cdc^{20}, a cofactor of APC/C and directly inhibits its function. The c-form of Mad2 can also bind to Mad1 which is immobilized on kinetochor (Sironi et al., 2002). According to the prion-like template model this Mad1-Mad2 complex is responsible for recruiting from the mitotic cytosol o-Mad2 and its transformation

to the c-form for Cdc^{20} binding (De Antoni et al., 2005; Mapelli et al., 2007; Mapelli and Musacchio, 2007; Yang et al., 2008).

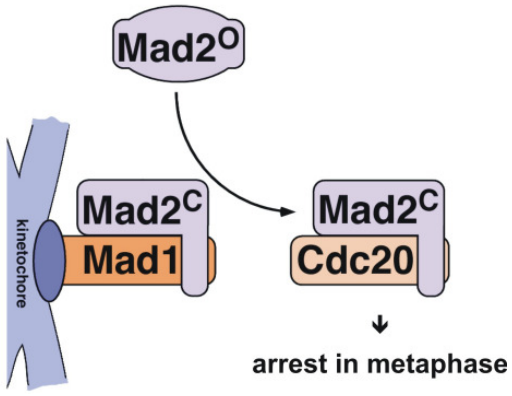

Figure 4.1.3. A prion like model of the spindle assembley checkpoint. Figure reproduced from Stemmann, (2008).

4.1.4 Shugoshin

The different regulation of sister chromatid cohesion at centromeres and along chromosome arms in mitosis is reminiscent of meiosis. Here, centromeric cohesion, but not arm cohesion, persists throughout anaphase of the first division. A protein required to protect centromeric cohesin Rec8 from separase cleavage has been identified and named shugoshin or Sgo1 (Kitajima et al., 2004). It has been established that shugoshin represents a conserved protein family defined as a centromeric protector of Rec8 cohesin complexes in meiosis, for instance it shows homology with Drosophila Mei-S332 and several uncharacterized proteins in other eukaryotic organisms (Rabitsch et al., 2004). During mitosis in vertebrates the cohesion is much more robust at the centromere at metaphase, where it antagonizes the pulling force of spindle microtubules and again shugoshin is essential to protect the centromeric cohesin, in this case Scc1, until late metaphase (Watanabe and Kitajima, 2005).

Currently in *S. pombe* as well as in humans (mammalians) two shugoshines Sgo1 and Sgo2 have been described, whereas in *S. cerevisiae*, *D. melanogaster* as well as in frogs and fishes only one has been found. They are highly divergent in sequence and length but all of them share a couple of structural features. They

posses a N-terminal coiled coil, conserved, about 30 amino acids long, C-terminal basic region and a connecting central segment rich in charged residues (Kitajima et al., 2004).

Although it is clear that shugoshin is capable of shielding cohesin from proteolityc action the exact mechanism regulating this process is still not explored. For example in meiosis shugoshin gets degraded after the first nuclear division but presence of shugoshin in meiosis II due to overexpression is not sufficient to inhibit separation of sister chromatids (Rabitsch et al., 2004). This indicates that shugoshin must be somehow inactivated after meiosis I as well as in late metaphase during mitosis. It is already known that shugoshin requires co-factors in order to protect cohesin. PP2A, a specific subtype of serine/threonine protein phosphatase 2A associates with human shugoshin and this way colocalizes at centromeres. Thus shugoshin is responsible for recruiting PP2A to the centromers. Purified Sgo1-PP2A complex has an ability to reverse the phosphorylation of cohesin in vitro, which indicates that the protection at the centromeres during mitosis is achived by dephosphorylation of Scc3 cohesin. Meiotic shugoshin (Sgo2) of fission yeast also recruits PP2A to the centromer where both proteins collaboratively protect Rec8-containing cohesin (Kitajima et al., 2006). At the same time the kinases Bub1 as well as aurora B maintain the steady-state levels and centromeric localization of shugoshins (Tang et al., 2004; Kitajima et al., 2005; Kitajima et al., 2006). However shugoshin not only localizes to the centromere but it was also found at kinetochores of checkpoint arrested vertebrate cells (Salic et al., 2004).

According to recent biochemical analysis *Xenopus laevis* shugoshin (xSgo1) is able to bind c-Mad2. The same was found for human shugoshin 2 (hSgo2) but not for human shugoshin 1 (hSgo1). This suggests a contribution of Mad2 by shielding cohesin from proteolityc action of separase (Stemmann, 2008).

4.1.5 Chromosome missegregation and cancer

Many mechanisms have been found to be responsible for cancer. Generally, an uncontrolled cell growth is possible through accumulation of mutations of proto-oncogenes or tumor suppressor genes and is a consequence of the cell cycle deregulation (more details in chapter 5.1.7). Also the contribution of genomic instability to tumor development is beyond dispute (Wang et al., 2002). The chromosomal instability (CIN) and microsatellite instability (MIN) were intensively investigated. Mitosis is the culmination of the cell cycle and it is also potentially dangerous when sister chromatids are irreversibly segregated to daughter cells. Defects in the checkpoints can lead to CIN and aneuploidy which are characteristic for many of human cancers (Jallepalli and Lengauer, 2001; Pihan and Doxsey, 2003). Whereas the connection between a MIN phenotype and cancer is now proven, the hypothesis that CIN causes cancer is still under debate. However, CIN clearly facilitates tumor development towards enhanced growth, metastatic potential, loss of apoptotic mechanisms and resistance to therapy (Marx, 2002). Nonetheless, the ubiquity of aneuploidy in human cancers, particularly solid tumors, suggests a fundamental link between errors in chromosome segregation and tumorigenesis (Ramaswamy et al., 2003; Draviam et al., 2004).

A compromised mitotic checkpoint has been reported for a number of cancers where the components including shugoshin, securin or APC/C are deregulated (Cahill et al., 1998; Wasch and Engelbert, 2005). Mutations in Mad1, Mad2, or Bub1 can result in premature separation of sister chromatids and, in succession, chromosome missegregation (Kops et al., 2005; Michel et al, 2001). On the other hand overexpression of shugoshin, found in 9 out of 10 breast cancers (Scanlanet al., 2001), as well as overexpression of securin (also referred to as pituitary tumor transforming gene, *PTTG*) and overexpression of Mad2 cause chromosome non-disjunction which leads to cancer due to non-reciprocal translocations and missegregations of chromosomes (Hamid and Kakar, 2003; Margolis, 2005).

4.2 Goals of the study

The acquisition of genomic instability is a crucial step in the development of human cancer. Chromosome missegregation caused by a compromised mitotic checkpoint has been reported for a number of tumors. Shugoshin plays a pivotal role in the process of chromosome segregation but the mechanism of its function is poorly understood. Also a lot of questions about its biochemical characterization remain still open. Recent studies (Stemmann, 2008) suggest that shugoshin directly interacts with the key checkpoint component Mad2 at kinetochores and this way both cooperate in shielding cohesin from proteolityc action of separase before anaphase begins. Overexpression of some mitotic checkpoint components like Mad2 or shugoshin is frequent in cancer cells and relates to metastasis and poor clinical outcome. For this reason a better understanding of shugoshin and its relation with Mad2 can be of great medical importance.

Structural analysis of Mad2 in complex with shugoshin could improve our knowledge of the mitotic checkpoint and its regulation. A precise model of the interaction between Mad2 and Sgo can be created based on high-resolution 3-D structures of the interacting proteins as well as kinetic data from binding studies by means of ITC or NMR experiments.

The main goal of the work presented in this part of the thesis was to study the molecular basis of relationship between Mad2 and shugoshin by means of x-ray crystallography. To support this research and enable characterization of Mad2 - Sgo interaction ITC and NMR analysis were employed.

The Mad2 – shugoshin interaction will also be used to illustrate a novel NMR-based drug screening method described in chapter 5.

4.3 Results

This section of the thesis contains results of the work on Mad2 and its binding partner shugoshin. The main goal of the project was the characterization of the Mad2 - Sgo interactions using various biochemical techniques, ITC, NMR spectroscopy and x-ray crystallography. The presented studies were carried out on several *Xenopus leavis* Mad2 constructs and *Xenopus leavis* as well as human shugoshin fragments of different length.

First methods of protein cloning, expression and purification are described. The folding analyses by NMR and the results of investigation of interaction between proteins are presented. Then the crystallization trials are reported.

4.3.1 Cloning, expression and purification of Mad2 and Sgo constructs.

4.3.1.1 Construct design and cloning

The construct design and choice of expression vectors are described in detail in Chapter 2. All constructs of Mad2 and Sgo were designed, expressed and purified based on previously published literature (Sironi et al., 2002). The gene of Mad2 was amplified from the *Xenopus laevis* oocyte cDNA (Stemmann, 2006). Mutation RQ133/134EA of the Mad2 was designed in order to secure the so called “close” conformation which binds Sgo (Nasmyth, 2005; Stemmann, 2006). The Mad2 binging site on Sgo, which is located between the N-terminal coiled-coil and the central segment (Figure 4.3.1), as well as the length of the Sgo fragments used for the presented experiments was determined by various methods, including limited proteolysis, mass spectroscopy and Far Western (Stemmann, 2006). Because of the high degree of homology of the Mad2-binding motif among other proteins interacting with Mad2 (Stemmann, 2006), some peptides used for MNR binding studies were designed on the base of Cdc20 and Mad1 sequences. The list of the peptides used in this work is shown in Table 4.3.1. The constructs of Mad2 and Sgo with references are listed in Table 4.3.2.

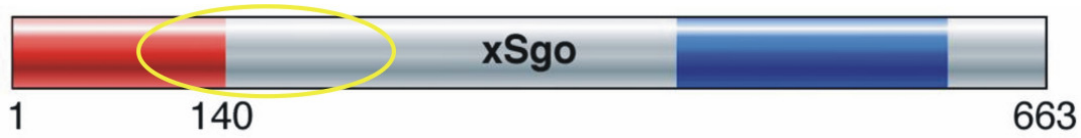

Figure 4.3.1. Architectural hallmarks of Shugoshin. N-terminal coiled coil shown in red, conserved C-terminal basic region in blue. Yellow marks the Mad2 binding region. The connecting, central segment is rich in charged residues. Figure modified from Stemmann, (2006).

Table 4.3.1. List of peptides used for the studies

	Name	Length (no. of aa)	Aminoacid seqence
1	**xSgo6**	6	$S^{166}AILRL^{171}$
2	**xSgo9**	9	$S^{166}AILRLPIH^{174}$
3	**xSgo25**	25	$S^{155}KVEEKLSSGASAILRLPIHAPSID^{179}$
4	**xMad1**	9	$T^{541}KVIHLSLN^{549}$
5	**xCdc20**	6	$A^{135}KILRL^{140}$
6	**hSgo16**	16	$K^{143}RISKQCLMRLPFAR^{158}$
7	**hSgo16P**	16	$K^{143}RIS(PO_4)KQCLMRLPFAR^{158}$

Table 4.3.2. List of Mad2 and Sgo constructs used in this thesis

	Construct name	Vector	Expression in *E. coli*
1	**His6-xMad2***	pQE80-FA	high, soluble, native
2	**xSgo(143-183)***	pMal-Tev2-FA	high, soluble, partially folded
3	**xSgo(1-207)***	pMal-Tev2-FA	high, soluble, unfolded
4	**xMad2Sgo(51-188)**	pRFSDuet-1	high, soluble, partially folded
5	**xMad2Sgo(105-188)**	pRFSDuet-1	high, soluble, partially folded

* constructs received from collaborators (Stemmann, 2006)

Properties of the designed constructs, such as their theoretical isoelectric points (pI), molecular masses (MW) and extinction coefficients (E_{280}) are shown

in Table 4.3.3. The information was deduced on the basis of the amino acid composition with the aid of the program ProtParam on www.expasy.org.

Table 4.3.3. Physicochemical properties of the investigated constructs

	Construct name	no.of aa	MW (Da)	E_{280} ($M^{-1}cm^{-1}$)	pI
1	**His6-xMad2**	209	24119	26510	5.83
2	**xSgo(143-183)**	41	4272	-	8.18
3	**xSgo(1-207)**	207	23371	1615	9.87
4	**xMad2Sgo(51-188)**	341	38566	28670	6.97
5	**xMad2Sgo(105-188)**	287	32217	27180	5.42

4.3.1.2 Expression and purification

All proteins were expressed in *E. coli* BL21(DE3) strain. Examination of the already existing methods of purification of Mad2 and similar or homologous proteins was used in designing the purification pathway. Purification of all used constructs and complexes followed a similar protocol, modified from Sironi et al., (2002). It employed two different liquid chromatography steps: affinity chromatography on Ni-NTA or amylose beads followed by gel filtration chromatography in the buffer proper for further applications of the protein. All expressed proteins or complexes were highly soluble and tagged which enabled the use of affinity chromatography under native conditions. Major variations of the procedure concerned the affinity chromatography and conditions of the different tag cleavages (determined by the use of different tags and enzymes). The strategy for purification of His6-xMad2-xSgo(143-183) and His6-xMad2-xSgo(1-207) complexes is shown in Figure 4.3.2. The general strategy for purification of His6-Mad2 as well as other xMad2-Sgo complexes is shown in Figure 4.3.3.

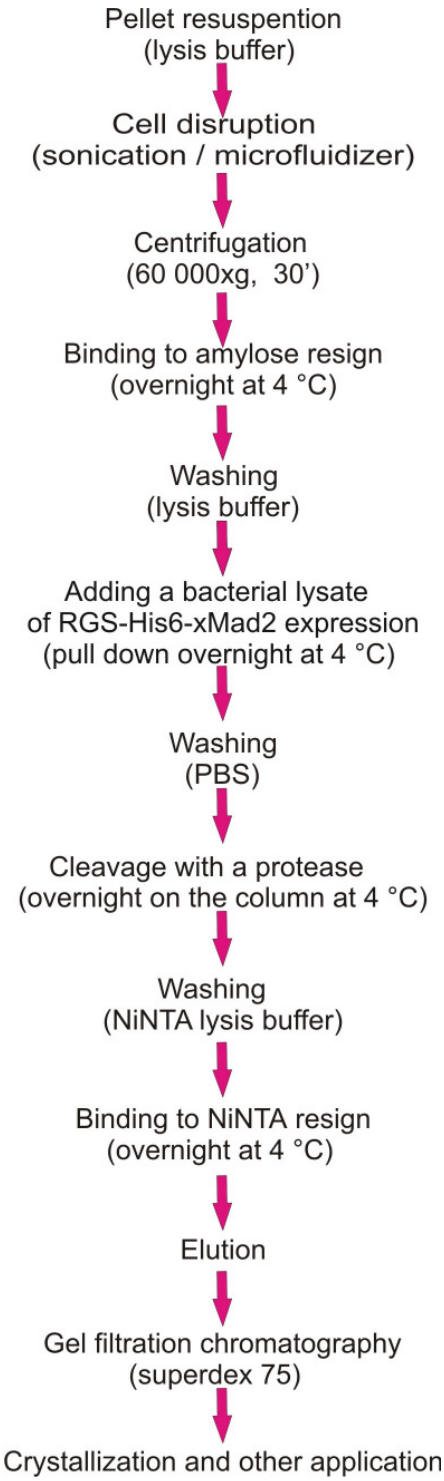

Figure 4.3.2. Flow chart of the purification schemes for His6-xMad2-xSgo co-purified complexes.

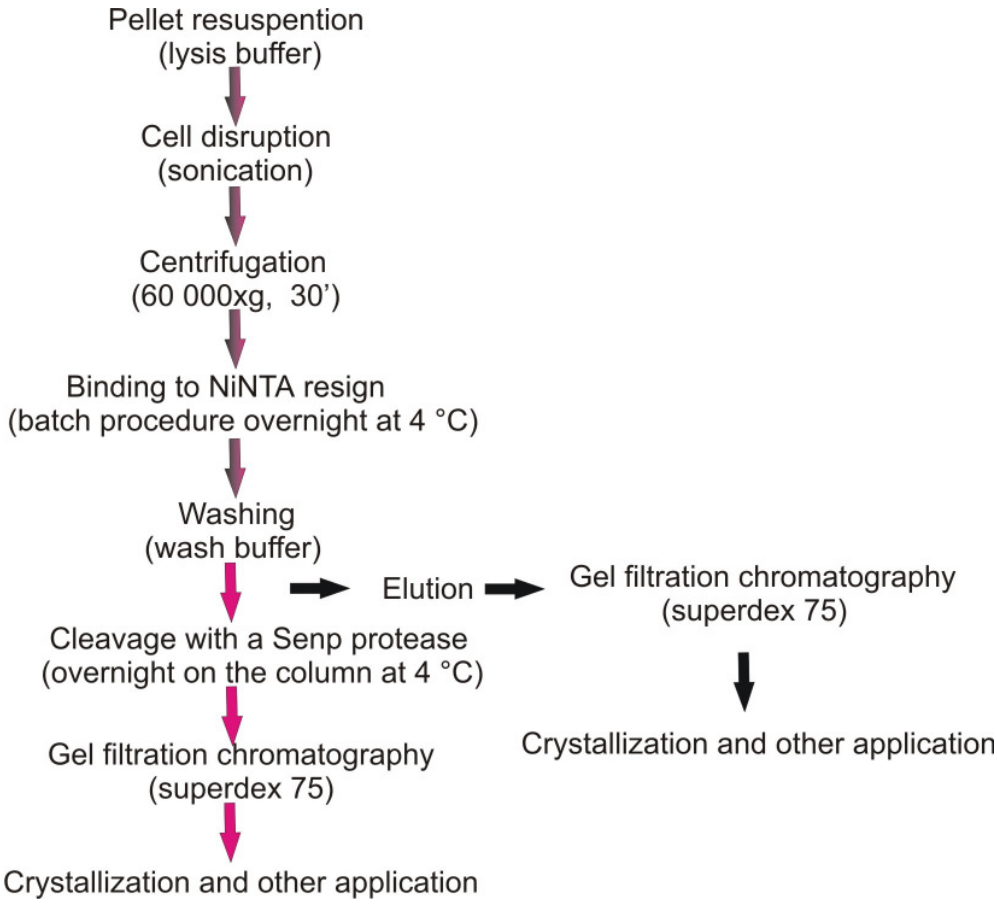

Figure 4.3.3. Flow chart of the purification schemes for His6-xMad2 (black arrows) and xMad2-xSgo coexpressed complexes (red arrows).

Gel filtration was the final step of purification enabling separation of the proteins from possible aggregates, digested tags, and the restriction protease used for the tag cleavage. Complete exchange of the buffer was also possible, removing any low molecular weight substances that could interfere with NMR measurements. The column used, a 120-ml Superdex 75 prep grade, is characterized by good resolution in the range 10-75 kDa. To make the best of its capability, low flow rates (0.8 ml/min) were used and samples not larger than 6 ml were loaded onto the column. A typical chromatogram and the SDS-PAGE analysis of His6-xMad2 and Mad2-Sgo complex purification are shown in Figure 4.3.4. After purification mass spectroscopy measurements were performed to confirm the identity of purified constructs.

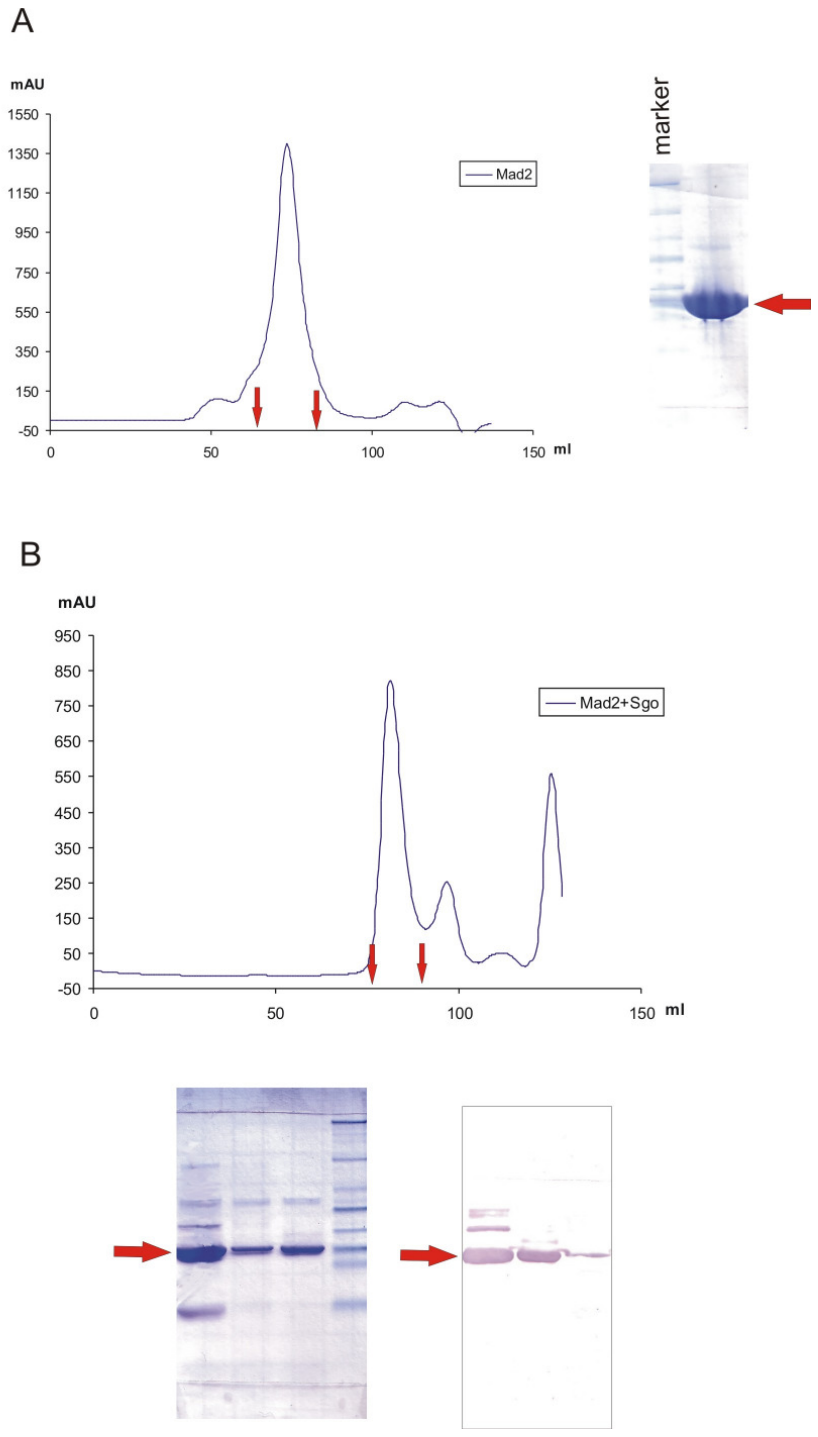

Figure 4.3.4. A typical chromatogram and the SDS-PAGE (with Western blot) analysis of the His6-xMad2 (A) and Mad2Sgo complex (B) after gel filtration. The fractions pooled are contained between the two arrows.

4.3.2 Functional and structural studies

4.3.2.1 NMR studies of the folding of Mad2 and Sgo

All investigated constructs were examined by NMR. 1D spectra were taken in order to prove the folding and structural integrity of purified protein. The spectra of His6-xMad2-xSgo(143-183) and His6-xMad2-xSgo(1-207) complexes exhibit a substantial peak at 8.3 ppm, suggesting that the complexes were only partially folded and contained large unstructured regions. The quality of the one-dimensional spectrum is dependent on the length of the construct. Spectra of His6-xMad2-xSgo(143-183) and His6-xMad2-xSgo(1-207) were very poor, most probably because of its large size and possible dimerization. For this reason spectra are not shown here.

Due to the poor features of NMR spectra and only partial folding of the His6-xMad2-xSgo co-purified complexes further structural studies, including crystallization trials, were not undertaken.

The His6-xMad2 construct gave spectra with a typical intensity pattern of a folded protein (Figure 4.3.5). The spectrum showed the signal dispersion downfield of 8.5 ppm and upfield of 1 ppm characteristic of structured proteins. The peak-width in this spectrum suggested that this construct might dimerize in solution.

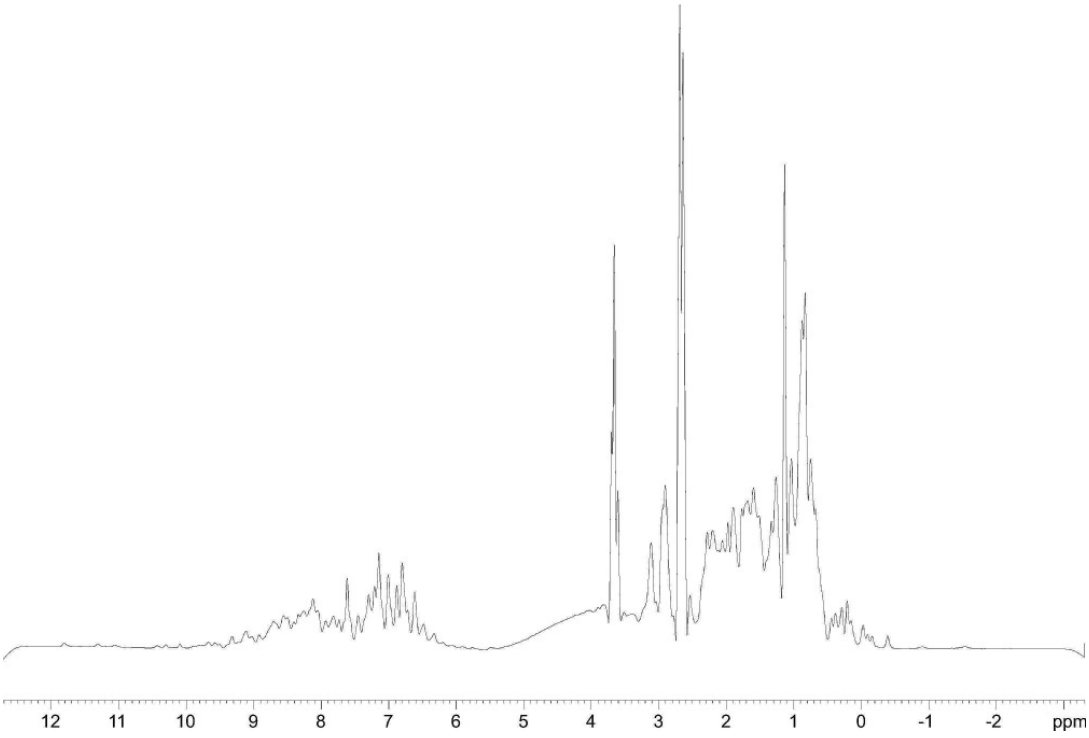

Figure 4.3.5. 1D proton NMR spectrum of the His6-xMad2 construct.

The chronologically latest constructs xMad2Sgo(51-188) and xMad2Sgo(105-188) were designed for the purpose of crystallization trials, after previous His6-xMad2-xSgo complexes turned out to be unfolded. The spectra show that both constructs are mostly structured (Figures 4.3.6 and 4.3.7). But based on these spectra it can be concluded that also a small part of each protein complex is in a random coil conformation.

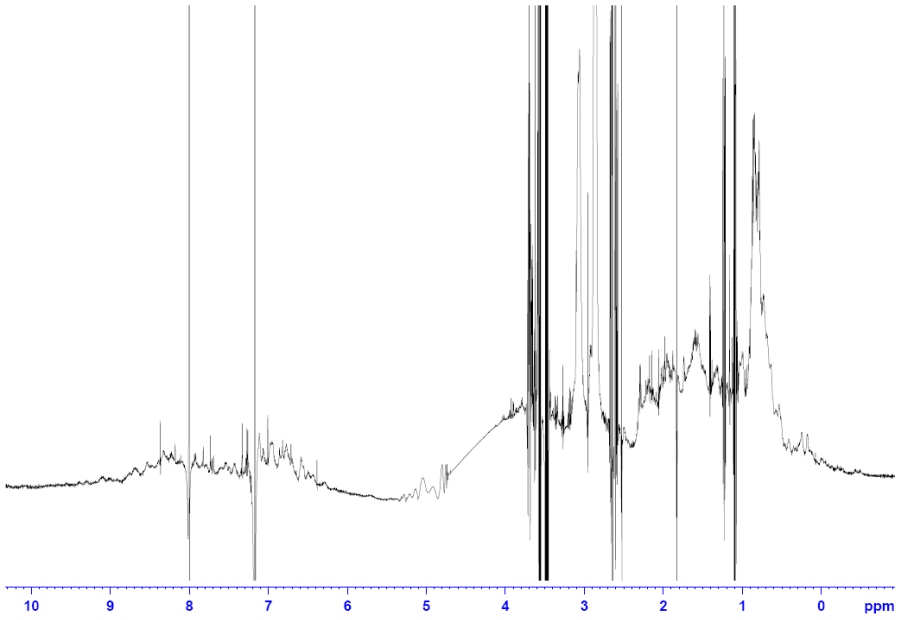

Figure 4.3.6. Characterization of the xMad2Sgo(51-188) construct by one-dimensional proton NMR spectrum.

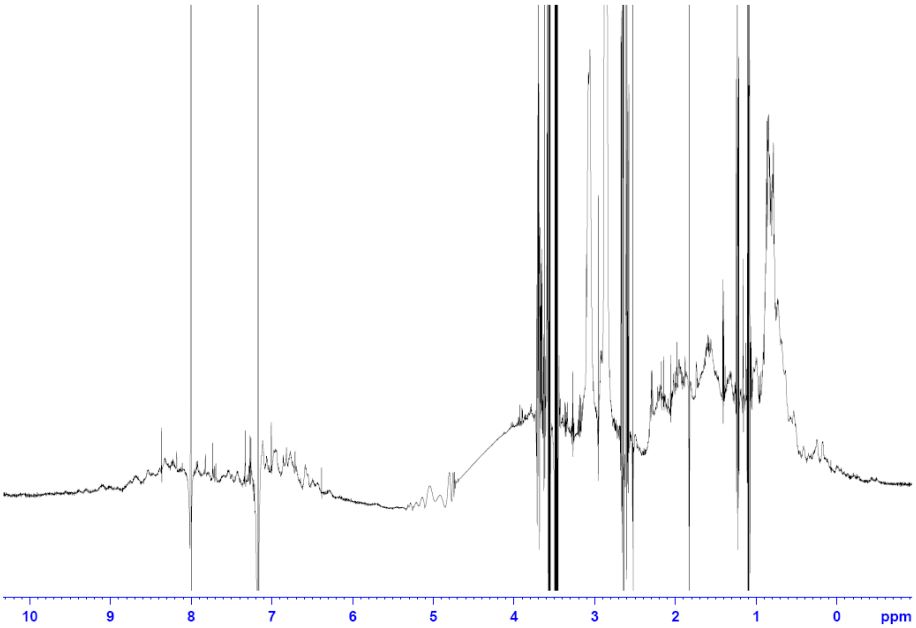

Figure 4.3.7. Characterization of the xMad2Sgo(105-188) construct by one-dimensional proton NMR spectrum.

4.3.2.2 NMR binding studies

The main goal of the binding studies was to understand the interaction between Sgo and Mad2.

In the first experiment His6-xMad2 was titrated with the xMad1 and xSgo9 peptides (for sequences see Table 4.3.1). Since the structure of the Mad1Mad2 complex is known (Sironi et al., 2002), this experiment should reveal whether both tested peptides bind to the same binding site on Mad2.

Both peptides were added in 4 steps to the His6-xMad2 protein solution leading to a molar ratio His6-xMad2: peptide of: 2.5:1; 1.25:1; 0.8:1 and 0.2:1. After each step a two dimensional HSQC spectrum was recorded.

Both peptides bind strongly (K_D<1µM) as seen by the splitting of the peaks during titration. In the middle step two signals from the starting and final position can be seen which indicates that complexes are in slow exchange. Also in both cases the same residues were affected which indicates that the xMad1 peptide as well as the xSgo9 peptide bind to the same binding site at His6-xMad2. The final positions of the signals for both peptides were different due to their different sequence and therefore different chemical environment. This does not indicate different binding sites (Figures 4.3.8 and 4.3.9).

The next experiment was carried out to compare the binding of hSgo16 and hSgo16P peptides to His6-xMad2 to check the influence of phosphorylation of ^{146}S on the binding features. The protein was titrated with 10 times excess of both peptides respectively. As well as in the previous titration with the xSgo9 and xMad1 peptides both peptides bind strong as seen by the splitting of the peaks during titration. Unfortunately the observed binding was too strong to conclude a potential difference in the binding K_D. Again in both cases the same residues were affected (Figures 4.3.10 and 4.3.11).

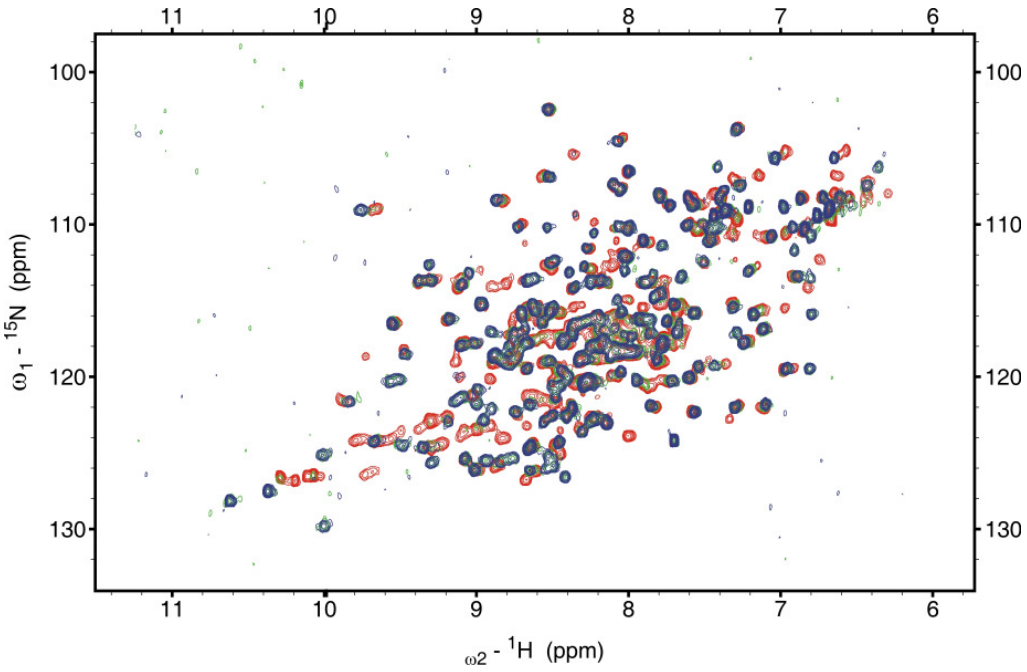

Figure 4.3.8. Titration of His6-xMad2 with xMad1 peptide. Red: reference spectra; Green: 1,25:1 and blue: the final step.

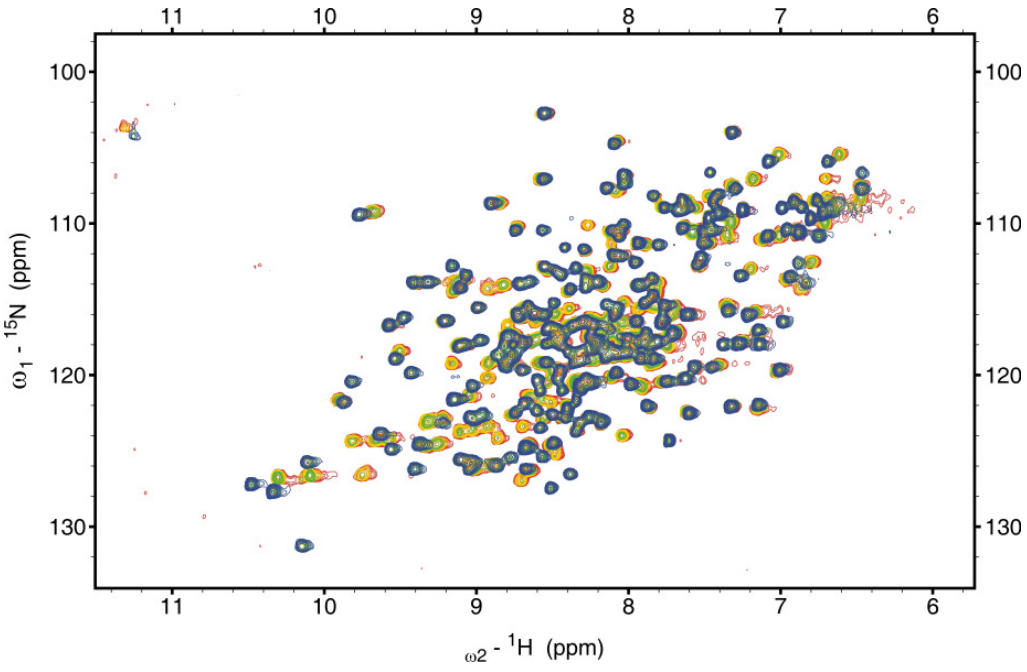

Figure 4.3.9. Titration of His6-xMad2 with xSgo9 peptide. Red: reference spectra; Yellow: 2.5:1; Green: 1,25:1 and blue: the final step.

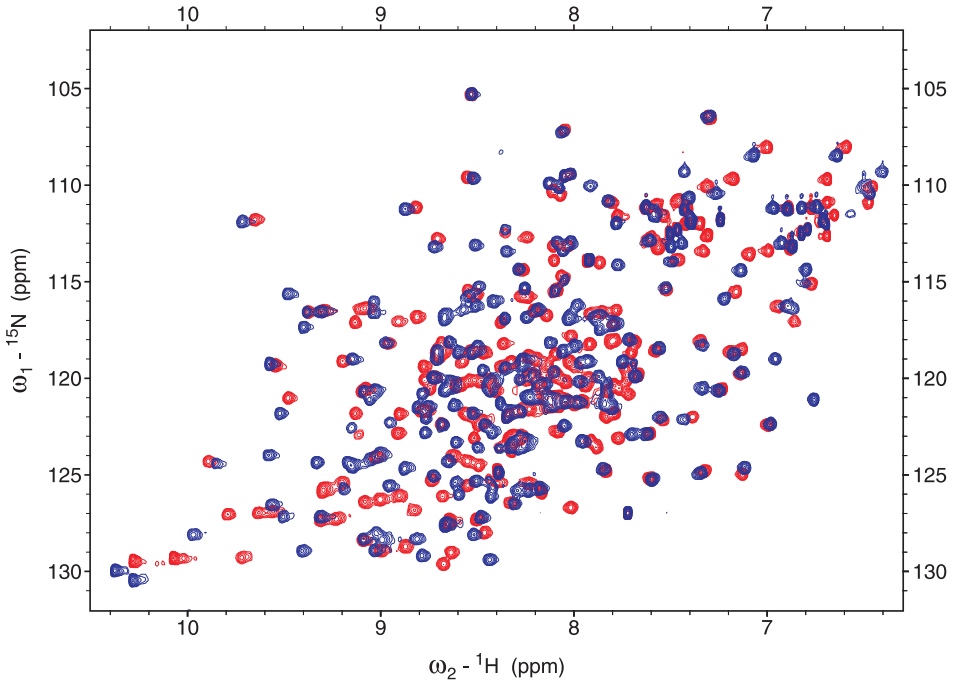

Figure 4.3.10. Titration of His6-xMad2 with the hSgo16 peptide. Red: reference spectra; blue: the final step.

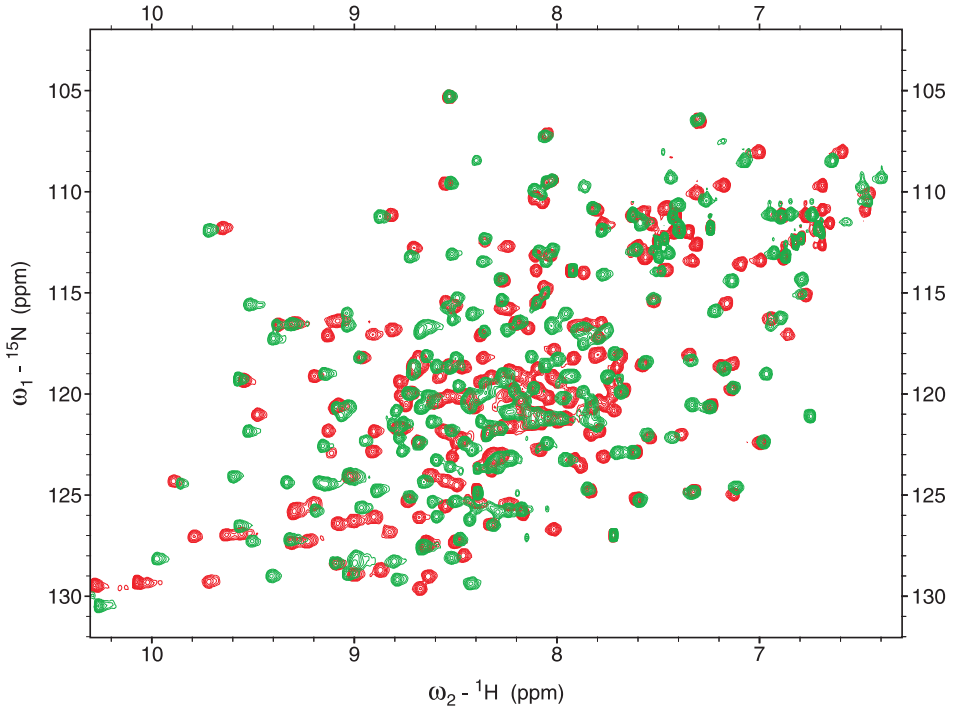

Figure 4.3.11. Titration of His6-xMad2 with hSgo16P peptide. Red: reference spectra; green: the final step.

This study of Mad2 and its binding partners illustrates a new approach to binding analysis by means of NMR (for details see Chapter 5). In some cases naturally present tryptophans might be used for detecting the binding. This is the case when tryptophans actively participate in the binding to ligands or are close enough to the binding pocket and are therefore sensitive in NMR experiments (Rothweiler et al., 2008). This can be seen in the titration of the protein His6-xMad2 with Mad1 or xCdc20 peptides (Figures 4.3.12 and 4.3.13).

The same method was employed to confirm the binding of the xSgo25 peptide to Mad2. The peptide was added in 5 steps to the His6-xMad2 protein solution leading to a molar ratio His6-xMad2: peptide of: 1:0.1; 1:0.25; 1:0.5; 1:1 and 1:5. Also in this case a stepwise addition of peptide leads to a splitting of the tryptophan peaks between 11.0 and 11.4 ppm (Figure 4.3.14), indicating strong binding.

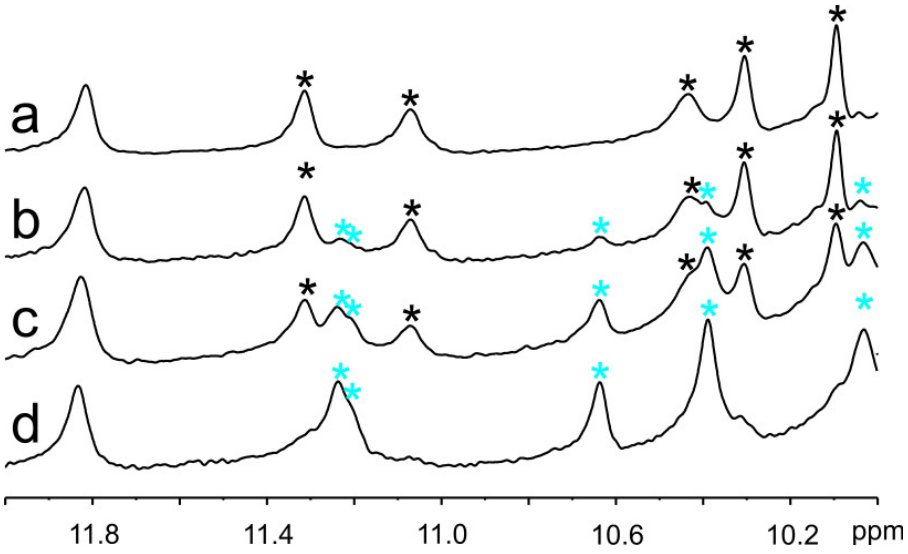

Figure 4.3.12. 1D NMR spectra of His6-xMad2 protein titrated with the xMad1 peptide. (a) Unbound His6-xMad2 (b) The molar ratio protein: peptide 1: 0.1. (c) Molar ratio protein: peptide 1: 0.4. (d) Molar ratio protein: peptide 1: 1. The peaks which are splitting are indicated with an asterisk (black: starting position, blue final position).

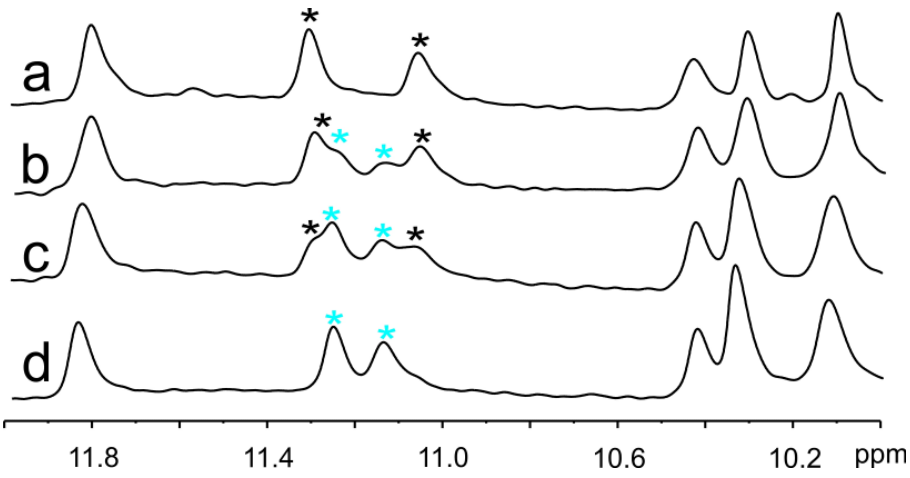

Figure 4.3.13. 1D NMR spectra of His6-xMad2 protein titrated with the xCdc20 peptide. (a) Unbound His6-xMad2. (b) The molar ratio protein: peptide 1: 0.3. (c) Molar ratio protein: peptide 1: 0.6. (d) Molar ratio protein: peptide 1: 1. The peaks which are splitting are indicated with an asterisk (black: starting position, blue final position).

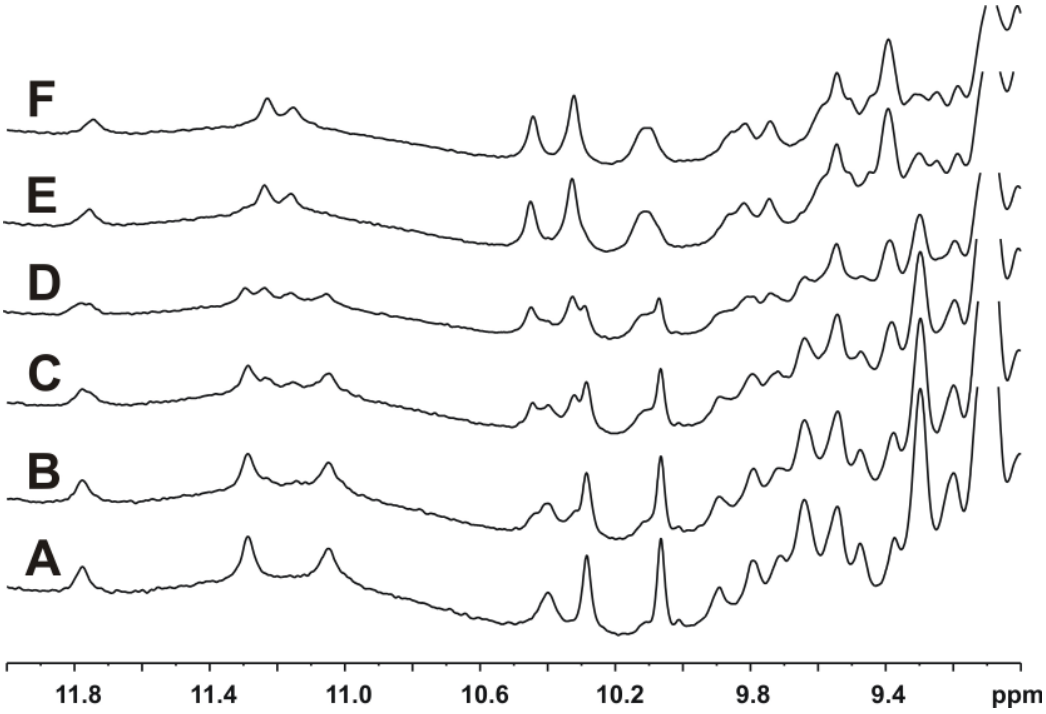

Figure 4.3.14. 1D NMR spectra of His6-xMad2 protein titrated with the xSgo25 peptide. (a) Unbound His6-xMad2. (b) The molar ratio protein: peptide 1: 0.1. (c) Molar ratio protein: peptide 1: 0.25. (d) Molar ratio protein: peptide 1: 0.5. (e) Molar ratio protein: peptide 1: 1. (f) Molar ratio protein: peptide 1: 5.

4.3.2.3 ITC measurements

ITC was used to study the interaction between His6-xMad2 and xSgo6, xSgo9 and xSgo25 peptides to compare how the length of the peptide influences the binding to the protein. ITC allows for precise measurement of K_D. Measurement of the heat generated during the interaction allows also determine thermodynamic parameters like dissociation constants or reaction stechiometry.

The experiment was carried out also for hSgo16 and hSgo16P peptides to check the influence of phosphorylation of Ser146 on the binding features. Table 4.3.4 summarizes the results of the ITC measurements. Experimental curves are shown below (Figures 4.3.15 and 4.3.16).

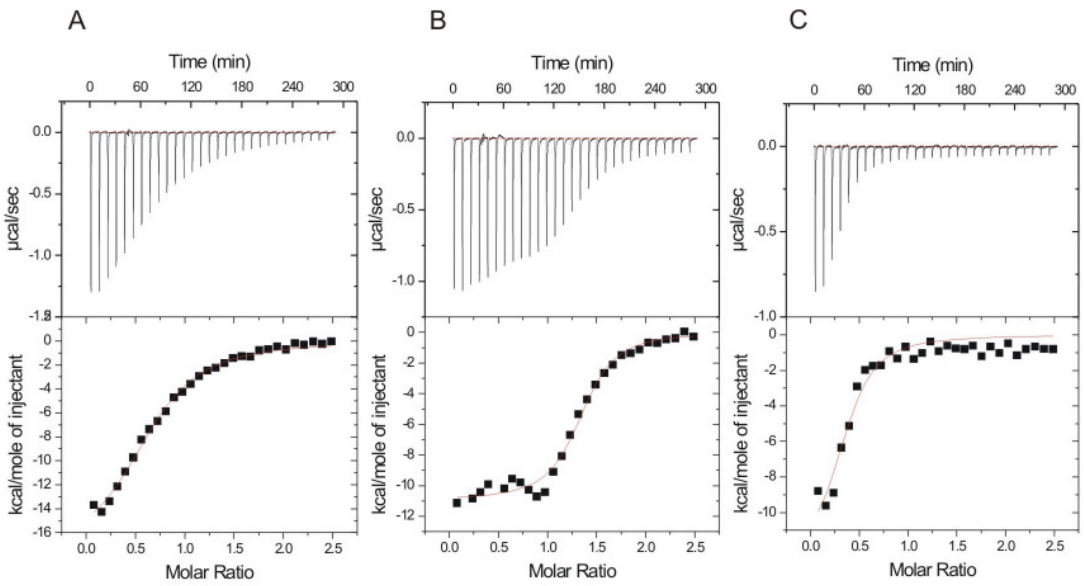

Figure 4.3.15. Examples of ITC data for the formation of the His6-xMad2 / xSgo complex. (A) His6-xMad2 titrated with xSgo6 peptide; (B) His6-xMad2 titrated with xSgo9 peptide; (C) His6-xMad2 titrated with xSgo25 peptide.

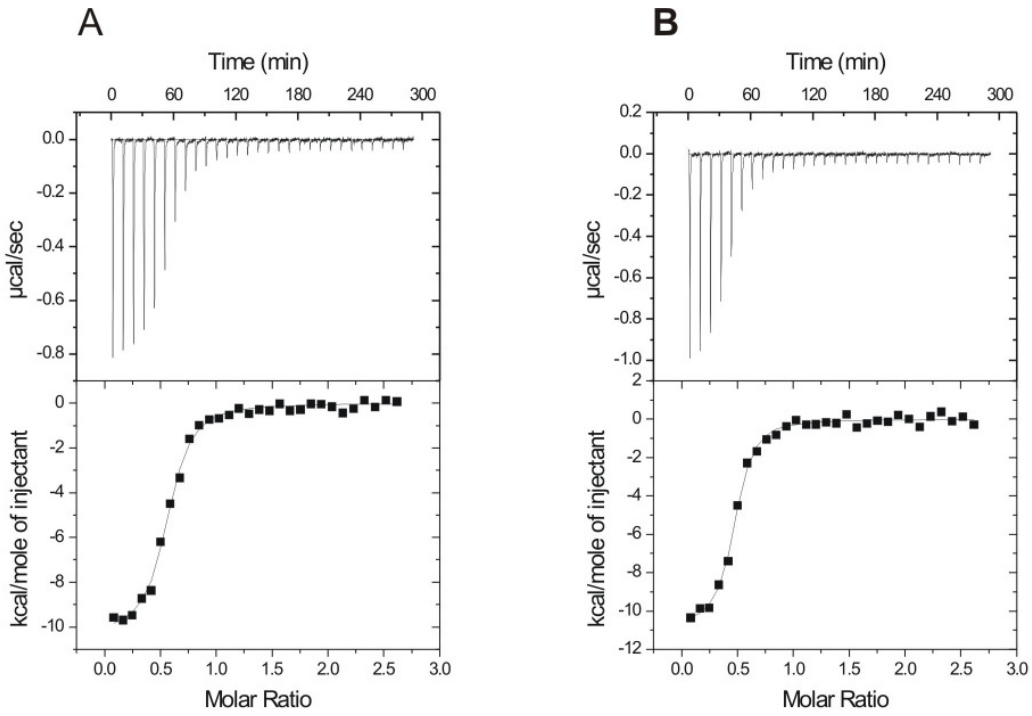

Figure 4.3.16. ITC data for the formation of the His6-xMad2 / hSgo16 complex. (A) His6-xMad2 titrated with hSgo16 peptide; (B) His6-xMad2 titrated with hSgo16P peptide.

Table 4.3.4. ITC data. Dissociation constants of diverse Mad2-Sgo complexes.

Complex	Reservoir	Titrant	K_D [µM]
His6-xMad2 / xSgo6	His6-xMad2	xSgo6	3,869
His6-xMad2 / xSgo9	His6-xMad2	xSgo9	0,604
His6-xMad2 / xSgo25	His6-xMad2	xSgo25	1,056
His6-xMad2 / hSgo16	His6-xMad2	hSgo16	0,685
His6-xMad2 / hSgo16P	His6-xMad2	hSgo16P	0,505

4.3.3 Crystallization

All the Mad2 constructs used for crystallization trials were checked by mass spectrometry, N-terminal sequencing, and NMR to confirm their purities, homogeneities and folding. Crystals were measured on a synchrotron source at the DESY facility in Hamburg. An extensive number of crystallization trials was

carried out during this thesis in order to get crystals of His6-xMad2, the His6-xMad2 / xSgo9 complex or later xMad2Sgo(51-188) and xMad2Sgo(105-188). Only in two of the large number of tested conditions for His6-xMad2 and His6-xMad2 / xSgo9 crystals appeared. Crystals of His6-xMad2 grew in Index Screen 36 and Grid Screen Ammonium Sulphate B2 (both Hampton Research) after a few days at 4°C (Table 4.3.5, Figure 4.3.17). Crystals from Index 36 were too small for measurements. Grid Screen Ammonium Sulphate B2 gave nice, needle-like, but unfortunately poorly diffracting crystals. Optimization of initial conditions did not improve the quality of the crystals.

Table 4.3.5. Crystallization conditions for His6-xMad2.

Crystallization condition	Reagent formulation
Index 36	15% v/v Tacsimate pH 7.0, 0.1 M HEPES pH 7.0, 2% w/v Polyethylene glycol 3,350
Grid Ammonium Sulfate B2	1.6 M Ammonium Sulfate, 0.1 M Citric acid pH 5.0

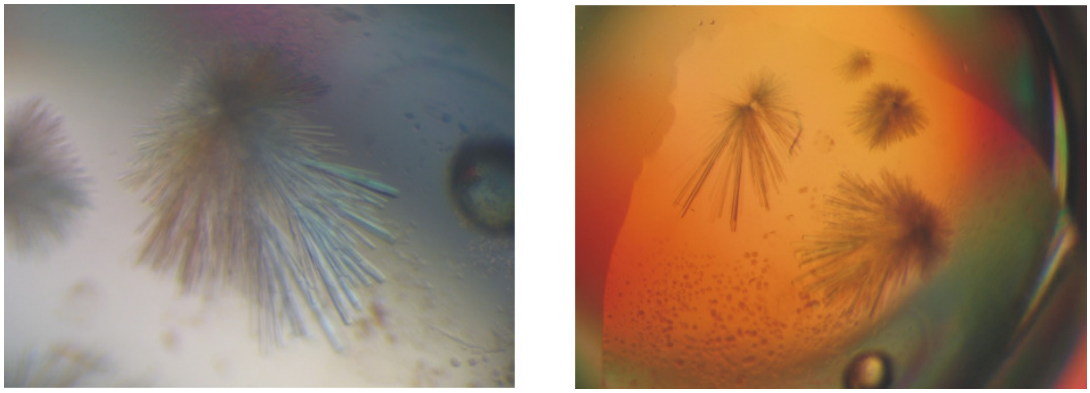

Figure 4.3.17. Crystals of His6-xMad2 grown in Grid Screen Ammonium Sulphate B2 at 4°C.

Crystallization of xMad2Sgo(51-188) and xMad2Sgo(105-188) was a very laborious task. More than 1600 crystallization conditions were unsuccessfully tested for each construct so far.

4.4 Discussion

Shugoshin is a central player in the process of cohesin protection from proteolityc action of separase during early phases of chromosome segregation but the exact mechanism as well as its biochemical characterization remains poorly understood. Several models are proposed for the physiological role of shugoshin. It is already known that shugoshin is responsible for recruiting PP2A to centromers. Sgo2-PP2A prevents separase-dependent cleavage of centromeric cohesin likely by dephosphorylation of Rec8 in meiosis I and Scc3 in mitosis (Kitajima et al., 2006). According to recent research (Stemmann, 2008), shugoshin binds directly to the key checkpoint component Mad2 at kinetochores.

Due to biochemical analysis it was shown that *Xenopus laevis* Shugoshin (xSgo1) is able to bind the closed conformation of Mad2. The same was proven for human shugoshin2 (hSgo2) but not human shugoshin1 (hSgo1). Moreover, xSgo1-Mad2closed can recruit Mad2open (Stemmann, 2008).

Mad2 in its closed form is immobilised through binding to Mad1 (Sironi et al., 2002), and the Mad1-Mad2 complex is responsible for recruiting open Mad2 from the mitotic cytosol and its transformation to the closed form for Cdc^{20} inhibition (Mapell and Musaccio, 2007; Yang et al., 2008). All this indicates a contribution of Mad2 by shielding cohesin from proteolytic action of separase.

The analysis of sequence alignment of Mad1, Cdc^{20} and shugoshin indicated that all of them share a conserved Mad2-binding motif (Figure 4.4.1). So far only the x-ray structure of the Mad1-Mad2 complex has been solved (Sironi et al., 2002).

Crystallization of the Mad2-Sgo complex could shed light on the structural basis for the mitotic checkpoint and its regulation, which was the main goal of the presented study.

Figure 4.4.1. Structural organization of the Mad2 binding region shown by way of Mad2 -Mad1 interaction. The sequence alignment indicates that Cdc[20], Mad1 and shugoshin share a conserved Mad2-binding motif. Figure adapted from Stemmann, (2008).

The initial idea of crystallizing the His6-xMad2-xSgo co-purified complexes was given up due to the poor features of the recorded NMR spectra, which suggested that the complexes were only partially folded and contained large unstructured fragments.

For a second approach a His6-xMad2 construct was chosen. Spectra of a 1-D NMR analysis showed the typical intensity pattern of a folded protein. If Mad2 could successfully be crystallized, it should be soaked with small fragments of Sgo. In case the binding between the protein and the Sgo peptide was strong, co-crystallization would have been tried. Unfortunately, crystals appeared only in two of the investigated crystallization conditions. In both cases measurements of the crystals did not yield satisfying results due to too small size or very poor diffraction. Optimization of initial conditions also did not improve the quality of the crystals.

Since the trials with Mad2 and a short Sgo peptide did not succeed, attempts to crystallize a longer complex were resumed. In this last attempt, the constructs xMad2Sgo(51-188) and xMad2Sgo(105-188) were designed. In this case Mad2 and shugoshin were co-expressed in the hope to achieve better folded and tighter bound complexes. Also in this construct a SUMO1 protein was co-expressed bound to Sgo. This step should improve shugoshin folding due to the excellent folding abilities of Sumo. During the purification the Mad2Sgo complex forms and Sumo is removed on the Ni column. The Sgo construct was

considerably shortened as compared to the initial version due to the NMR results of the previous constructs which had shown unfolded parts. The lengths of the Sgo constructs were chosen to encompass the binding site and predicted structural features. While NMR showed promising results, the subsequent crystallization trials were unsuccessful. The reason for this failure was most probably incomplete folding of parts of the complexes.

Recently Mapelli et al. (2007) reported the crystal structure of a Mad2 conformational dimer (Mad2open – Mad2closed). They speculate that the dimerization is highly flexible and for this reason some flexible residues were deleted from Mad2open and Mad2closed was stabilized by a binding peptide in the binding site which is also under investigation here. In future studies their results could be used to solve the structure of Mad2 in complex with shugoshin.

As the consecutive crystallization attempts turned to fail, NMR binding studies were employed to gain some insights into the Mad2-Sgo interaction. Structural as well as binding properties of proteins can be investigated by two-dimensional heteronuclear single-quantum coherence experiments (HSQC). The HSQC shows one NMR peak for every proton bound directly to a nitrogen atom and thus exactly one signal per residue in the protein (Rehm et al., 2002). Monitoring changes in chemical shifts and line widths of amide resonances of a ^{15}N-labeled protein upon addition of the binding partner allows to detect and to characterize possible protein-ligand binding. Additionally, the novel NMR-based drug screening method developed by our group and described in details in chapter 5 was also applied.

To bear out the hypothesis provided by our collaborators (Stemmann, 2006) that shugoshin binds Mad2 just like Mad1 and Cdc20 do, His6-xMad2 was titrated with the xMad1 and xSgo9 peptides. Both peptides bound strong (K_D<1µM) and in both cases the same residues were affected. This clearly indicates that the xMad1 peptide as well as the xSgo9 peptide bind to the identical binding site at His6-xMad2.

It can thus be concluded that Sgo, Mad1 and Cdc20 all bind Mad2 at the same binding site of the so called *safty belt* (Sironi et al. 2002).

Another HSQC experiment was carried out to check the influence of phosphorylation of ^{146}S of human Sgo on its binding features. His6-xMad2 was titrated with hSgo16 and hSgo16P peptides respectively. Also here, as in the case of the previous titration with xSgo9 and xMad1 peptides, both peptides bound strongly to Mad2 and the same residues were affected. The observed binding was too strong to conclude a potential difference in the binding K_D.

Because NMR studies could not give any exact information about binding strength, to support this research ITC experiments were carried out. At the beginning ITC was employed to determine the optimal length of the Sgo peptide for further experiments including crystallization. Out of three designed peptides xSgo9, a 9-amino acid long peptide containing the suspected Mad2-binding motive turned out to bind Mad2 with the lowest K_D in the nanomolar range (604 nM). The shorter one (6 amino acids) as well as the longer one (25 amino acids) bound 4- and 2-times weaker, respectively. This might indicate that the residues in the direct neighbourhood of the established binding motive have also influence on the binding between the proteins. On the other hand the fact that xSgo25 bound weaker than xSgo9 may result from spherical hindrance of potentially unfolded, larger fragment of the peptide.

ITC was chosen also to determine the potential difference in the binding strength between phosphorylated and not-phosphorylated human Sgo fragments to Mad2. The difference in obtained K_Ds was practically in the range of the measurement error which suggest that phosphorylation of ^{146}S of human Sgo is not of great importance by Mad2 binding.

The binding studies carried out by means of NMR and ITC confirmed the hypothesis that Shugoshin binds Mad2 the same way as Mad1 and Cdc20 do. This suggests that Mad1 as well as Cdc20 compete with shugoshin for binding to Mad2. *In vivo* Mad1 is required to form the Cdc20-Mad2 complex, Mad2 is partially immobilized at the kinetochore through binding with Mad1 and on the other hand should bind shugoshin to the same binding site. It comes as no surprise that no biochemical activity assay for the mitotic checkpoint has been described to date and several models can be evoked for the physiological role of

the shugoshin-Mad2 complex formation. Shugoshin might be activated by picking-up activated Mad2 at unattached kinetochores, the shugoshin-Mad2 complex might then diffuse to nearby centromeres to protect cohesin. Concurrently, as PP2A and Mad2 simultaneously interact with shugoshin, and this binding leaves PP2A unaffected, Mad2 may act as a co-protector of cohesin by keeping Sgo2-PP2A at centromeres until chromosomes have properly attached (Stemmann, 2008).

5 CDK2

5.1 Biological background

In the 1950's E. Krebs and E. Fisher discovered protein phosphorylation and protein kinases (Krebs and Fisher, 1955; 1956; 1958). This was the beginning of an amazing advancement in the understanding how protein kinases are regulated and control cell fate. Among the estimated 1000 human protein kinases the family of cyclin dependent kinases (CDKs) has been extensively studied because of their essential cellular functions (Cohen, 2002; Knockaert et al., 2002). CDKs play a pivotal role in controlling the cell cycle (CDK1-7), in apoptosis (CDK2), in neuronal function (CDK5) and in transcription (CDK7, 8, 9).

5.1.1 The cell cycle

The cell division is the essential mechanism by which all living beings reproduce. Its most basic function is to duplicate accurately the vast amount of DNA in the chromosomes and then segregate the copies precisely into two genetically identical daughter cells. These processes define the two major phases of the cell cycle. The details vary from organism to organism and at different times in an organism's life. Certain characteristics, however, are universal (for a comprehensive overview see Alberts et al., 2002).

In the original concept, the cell cycle was divided into two stages: mitosis (M), i.e. the process of nuclear division; and the interphase. Cells in the interphase seem to simply grow in size but more advanced research verified that the interphase includes G1, S and G2 phases (Norbury and Nurse, 1992). Thus, the eukaryotic cell cycle is traditionally divided into four sequential phases: G1, S, G2 and M (Figure 5.1.1). Replication of DNA occurs in a specific part of the interphase called synthesis (S) phase. The S phase is preceded by a brake called G1 during which the cell is preparing for DNA synthesis and is followed by a brake called G2 when the cell prepares for mitosis (Vermeulen at al., 2003). These gap phases allow more time for the growth and doubling of the mass of

proteins and organelles. Moreover they provide time to monitor the internal and external environment to ensure that conditions are suitable and preparations are complete before the cell commits itself to division (Alberts et al., 2002).

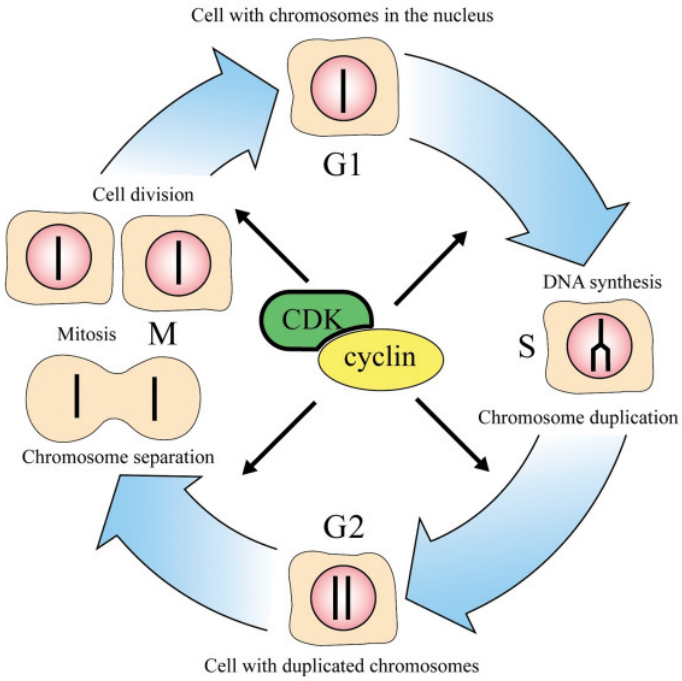

Figure 5.1.1. The different phases of the cell cycle. The cell grows continuously in the interphase (G1-S-G2). In the M phase, the nucleus and then the cytoplasm divide. Figure adapted from www.nobelprize.org.

5.1.2 CDKs and the control of the cell cycle

For a long time the regulation of the cell cycle was a *terra incognita* inside the cell. The major breakthrough came in the late 1980s with the identification of the key proteins of the cell cycle control system (for a comprehensive overview, see Alberts et al., 2002).

The transition from one cell cycle phase to another occurs in an orderly fashion and is regulated by different cellular proteins. At the heart of the cell cycle control system is a family of the protein kinases known as cyclin-dependent kinases (CDKs).

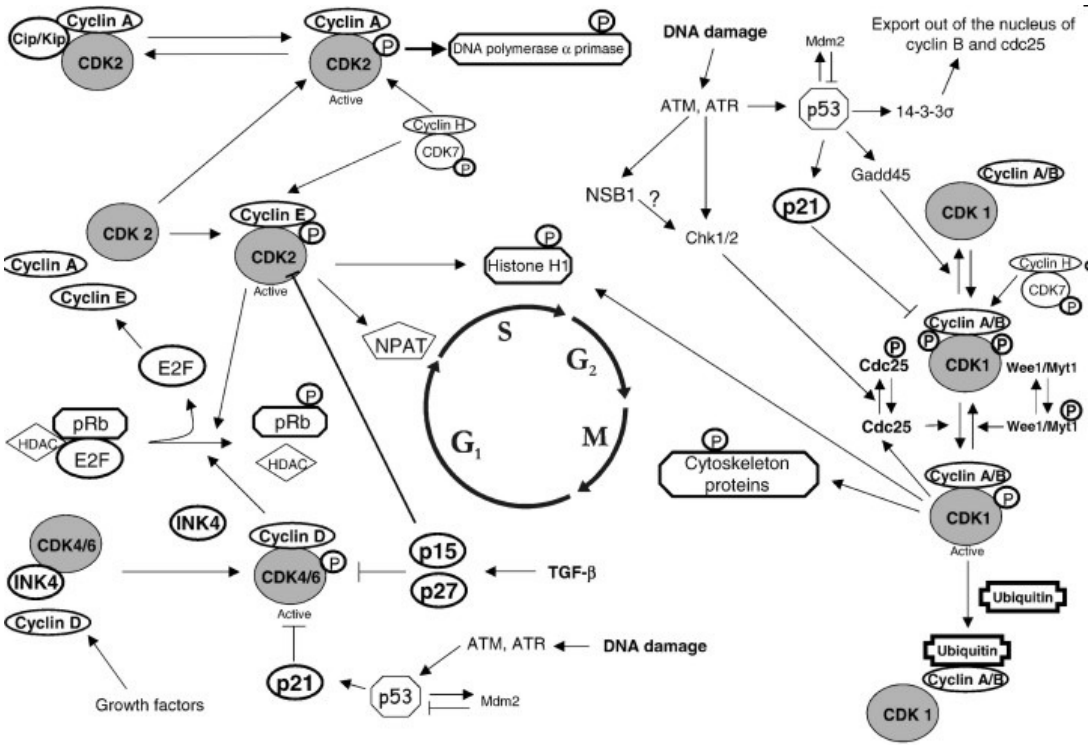

Figure 5.1.2. Schematic overview of some essential steps in the cell cycle regulation. Figure reproduced from Vermeulen et al., (2003). Details regarding the different phases are described in the text.

The transient activation of various CDKs controls the progression of the cell division through the G1→S→G2→M phases (Figure 5.1.2). When activated, CDKs induce downstream processes by phosphorylating selected proteins (Morgan, 1995). CDK protein levels remain stable during the cell cycle, in contrast to their activated proteins, the cyclins. Cyclin protein levels rise and fall as the cell progresses through the cycle, which was also the reason for naming these proteins “cyclins” after their discovery by Hunt in the early 1980s. Cyclical changes in cyclin levels result in the cyclic assembly and activation of the cyclin-CDK complexes while this activation in turn triggers cell cycle events (Evans et al., 1983; Pines, 1991).

CDK2 was the first member of the CDK family, which was discovered in the cell cycle of yeast. At present, thirteen CDKs and 25 cyclin box-containing proteins have been identified in humans (Knockaert et al., 2002). Those known to

interact in a complex are listed in Table 5.1.1. In addition there are several CDK related kinases with no identified cyclin partner. These are recognized by their sequence homology and by the presence of a variation of the conserved motif '*PSTAIRE*' in CDKs located in the cyclin binding domain. The variations are for example: *PCTAIRE*, *PFTAIRE*, *PITAIRE*, *KKIALRE*, *PISSLRE* and *PITSLRE*. Cyclins F, G, M, O, and P do not have any CDK partners and are known as 'orphan' cyclins (Meijer et al., 2000; Meyerson et al., 1992).

Table 5.1.1. CDKs known to pair with cyclins and identified for their distinctive role in the cell cycle (Meijer et al. 2000)

CDK	Partner	Function
CDK1	cyclin B	Prophase to metaphase transition, regulation of topoisomerase-2
	Ringo	β-Amyloid-induced cytotoxicity
CDK2	cyclin A	G1→S transition, S phase and G2 phase, centrosome duplication, regulation of Sp1-mediated transcription
	cyclin E	Apoptosis in thymocytes, mesangial cells, DNA damage-induced apoptosis
CDK3	cyclin E	G1→S transition
CDK4	cyclin D	G1 phase excitotoxin-induced neuronal cell death, dopamine and glutamate signaling, neurite outgrowth, neuron migration, induction of acetylcholine receptors
CDK5	p35, p25	Apoptosis
	p39, p29	Golgi membrane traffic, insulin exocytosis by β-cells phototransduction, neurotransmitter release (VDCC), Lens cell differentiation, myogenesis
CDK6	cyclin D	G1 phase neuronal cell death
CDK7	cyclin H	CDK1, 2 activation, basal transcription (TFIIH)
CDK8	cyclin C	Basal transcription, regulation of CDK7–cyclin H
CDK9	cyclin K	Signal transduction RNA transcription (P-TEFb), HIV-Tat-dependent transcription
	cyclin T	MyoD-mediated myocyte differentiation
CDK11	cyclin L	RNA processing or transcriptional apoptosis, dopamine and glutamate signaling

Different cyclins are required at different phases of the cell cycle. In early to mid G1 the first set of cyclins (D1, D2, D3) are activated by the mitogenic signals (Matsushime et al., 1991; Xiong et al., 1991; Motokura et al., 1991). Unlike the other cyclins, cyclin D is not expressed periodically but is synthesized as long as growth factor stimulation persists (Assioan and Zhu, 1997). Inhibition of this stimulation signaling pathway leads to cyclin D1 phosphorylation, enhances its nuclear export, and accelerates its ubiquitin-dependent proteasomal degradation in the cytoplasm (Hunter et al., 1994). G1 cyclins assemble with CDK4 and CDK6, which is essential for entry into the G1 phase (Sherr, 1994). CDK4 and CDK6 in turn get activated by phosphorylation, with a CDK activating kinase (CAK) (Matsushime et al., 1992; Meyerson et al., 1994). These activated CDK-cyclin D complexes (CDK4-cyclin D1 and CDK6-cyclin D3) phosphorylate the retinoblastoma protein (pRb). Phosphorylation of pRb causes dissociation of the E2F/DP1/pRb complex and turns off its transcriptional suppressor activity (de-Gregori et al., 1995; Arata et al., 2000; Sherr, 1996). Release of E2F/DP1 results in transcriptional activation of a set of genes required for entry into the S-phase including thymidylate synthetase, dihydro-folate reductase, cyclin A and cyclin E. Cyclin E, another G1 cyclin, activates CDK2, which acts to further hyperphosphorylate pRb and results in an irreversible commitment to cell division and transition into S-phase (Ohtsubo et al., 1995; Lukas et al., 1997; Lundberg et al., 1998).

During the S-phase, CDK2-cyclin A phosphorylates different substrates allowing DNA replication and inactivation of the G1 transcription factors (Giradr et al., 1991). In the late S phase CDK1 complexes with cyclin A. After that CDK1-cyclin B is formed and triggers the G2→M transition by phosphorylating substrates like the nuclear lamins. The same complex regulates further mitosis (King et al., 1994). Cyclin B proteolysis, transition to anaphase and completion of mitosis requires phosphorylation of the ‘Anaphase Promoting Complex’ (APC) by CDK1-cyclin B (King et al., 1996).

Some CDKs directly influence transcription. The CDK7-cyclin H/MAT1 complex is responsible for the phosphorylation of the C-terminal domain of the

large subunit of RNA polymerase II, which is required for the elongation process (Adamczewski et al., 1996; Feaver et al., 1994; Roy et al., 1994; Serizawa et al., 1995). Also, CDK8-cyclin C phosphorylates CTD RNA pol II, but on different sites.

5.1.3 Quality control of the cell cycle

Activation and inactivation of these successive waves of cyclin-CDK assemblies are tightly regulated by posttranslational modifications and intracellular translocations. They are coordinated through the so-called restriction point and checkpoint controls ensuring an orderly sequence of events in the cell cycle. If some malfunction prevents the successful completion of a previous step, signals are sent to the control system to delay progression to the next phase. The restriction point is defined as a point of no return in G1, following which the cell is committed to enter the cell cycle. Experiments demonstrate that cells starved of serum before the restriction point enter a G0-like state, while cells starved after restriction point are unaffected and continue through mitosis (Pardee, 1974). Entry into mitosis is prevented, for example, when DNA replication is not complete and chromosome separation in mitosis is delayed, if some chromosomes are not properly attached to the mitotic spindle (spindle checkpoint, described additionally in section 4.1). Progression through G1 and G2 is delayed by braking mechanisms, if the DNA in the chromosomes is damaged by radiation or chemicals. These delays provide time for the damaged DNA to be repaired, after which the cell cycle brakes are released and progress resumes. DNA damage checkpoints are positioned before the cell enters S phase (G1-S checkpoint) or after DNA replication (G2-M checkpoint) and there appears to be DNA damage checkpoints during S and M phases.

Checkpoints generally operate through negative intracellular signals (Alberts et al., 2002). The so-called ‘pocket protein’ family of cell cycle regulators pRb and the p107/p130 complex are required for the repression of distinct sets of genes potentially due to their selective interactions with E2Fs that are engaged at specific promoter elements (Aslanian et al., 2004; Cam et al., 2003; Classon et

al., 2001; Ginsberg 2004; Stevaux et al., 2002). In the G1 to S phase transition pRb binds to the E2F family of transcription factors and the resulting pRb-E2F complex not only blocks transcriptional activation by E2F but also forms an active transcriptional repressor complex at the promoter of cell cycle genes that can block transcription by recruiting histone deacetylase (HDAC) and the remodeling chromatin complex (SWI/SNF complex) (Rayman et al., 2002; Nicolas et al., 2003; Zhang et al., 2000; Harding et al., 2004). SWI/SNF proteins form a 11-subunit complex and when bound to nucleosomes alter the path of DNA bending around the nucleosome core in an ATP-dependent reaction that does not by itself displace the underlying histones. This complex is able to disrupt the nucleosome structure in a sequence–independent manner (Strober et al., 1996; Dunaief et al., 1994). The pRb-E2F-HDAC repression can be removed by the sequential phosphorylation of pRb that inhibits HDAC binding, thereby blocking active transcriptional repression (Hinds et al., 1992; Dowdy et al., 1993; Ewen et al., 1994; Kato et al., 1993; Hatakeyama et al., 1994).

At the G1-S checkpoint, cell cycle arrest induced by DNA damage is p53-dependent. Usually, the cellular level of p53 is low but DNA damage can lead to rapid induction of p53 activity (Levine, 1997). In this process different protein kinases are involved e.g. ataxia-telangiectasia-mutated (ATM), ataxia and rad3 related (ATR) kinases. They recognize DNA damage and in respons to it phosphorylate p53. p53 stimulates the transcription of different genes including p21, Mdm2 and Bax (Agarwal et al., 1998). The induction of p21 results in CDK inhibition and cell cycle arrest preventing the replication of damaged DNA (Ko and Prives, 1996).

When DNA damage occurs during G2, cells are able to initiate a cell cycle arrest in the presence or absence of p53. The entry into mitosis is prevented by maintaining CDK1 in its inhibited form through inhibitory phosphorylation or by sequestration of components of the CDK1-cyclin B complex outside the nucleus. This is achieved by the protein kinases Chk1 and Chk2, which are activated during DNA damage in an ATM-dependent manner and which phosphorylate Cdc25 (Vermeulen et al., 2003). Phosphorylation of Cdc25 inhibits its activity and

promotes its binding to 14-3-3 proteins, sequestering it outside the nucleus and preventing it from activating CDK1-cyclin B and mitotic entry (Sanchez et al., 1997; Zeng et al., 1998). Besides induction of inhibitory phosphorylations on CDK, p53 may also play a role in the regulation of the G2-M checkpoint. DNA damage-dependent increase of p53 results, as during the G1-S checkpoint, in increased transcription of p21 and of 14-3-3 sigma (14-3-3 σ). Increased binding of cyclin B to 14-3-3 σ actively excludes it from the nucleus. p53 also mediates the dissociation of CDK1-cyclin B1 complexes by induction of Gadd45 (growth arrest and DNA damage inducible gene) (Hermeking et al., 1997; Taylor and Stark 2001).

5.1.4 The regulation of CDKs

CDKs are composed of a protein kinase catalytic subunit and a regulatory cyclin subunit. As their name implies CDKs are dependend on cyclins for their activity. The typical CDK catalytic subunit contains a 300 amino acid catalytic core that is completely inactive when monomeric and unphosphorylated (De Bondt et al., 1993; Morgan et al., 1994). The inactive monomeric CDK subunit becomes activated through protein conformational changes accompanied by its association with the specific cyclin molecule and phosphorylation of a conserved and essential threonine residue within the activation segment of the kinase (Thr160 of CDK2) (Kobayashi et al., 1992; Morgan et al., 1994).

The control of activity of CDKs is complex. Each CDK interacts with a specific subset of cyclins. Cyclin function is primarily controlled by changes in cyclin levels, which increase characteristically at specific cell cycle stages and with response to the cell signals. Homology among cyclins is often limited to a relatively conserved domain of about 100 amino acids, the cyclin box, which is responsible for CDK binding and activation (Kobayashi et al., 1992; Lees et al., 1993). Constant synthesis during the cell cycle results in a roughly linear increase in cyclin concentration until at mitosis cyclin degradation abruptly increases, resulting in a rapid decline in cyclin levels.

Final activation of the CDKs requires dephosphorylation of the inhibitory pThr14 and pTyr15 sites by the dual specificity phosphatase CDC25. CDC25 activity decreases during mitosis, largely because of increased phosphorylation. During mitosis, the kinase responsible for this phosphorylation is activated and the corresponding phosphatase is inhibited (Zeng et al., 1998; Gyuris et al., 1993; Harper et al., 1993; Hanon et al., 1994). The kinases responsible for these two phosphorylations are the Myt1 and Wee1 kinases, which phosphorylate CDK1 on Thr14 and Tyr15, respectively (Morgan et al., 1994). Genetic and biochemical studies have implicated a second phosphorylation event important for CDK regulation. KAP (kinase-associated phosphatase also termed Cdi1, Cip2) as one of the protein phosphatases seems to be responsible for dephosphorylating the pThr160 residue of human CDK1, CDK2 and CDK3 which results in their activation.

The CDK inhibitors also known as CKIs play an important role in CDK regulation. They can either physically block activation or block substrate/ATP access. Most antiproliferative signals lead to the induction of CDK inhibitors. For example senescence, contact inhibition, extracellular anti-mitogenic factors like TGF-β and cell cycle check points like the p53 DNA damage, induce p16INK4a, p27Cip2, p15INK4b, and p21Cip1, respectively. The inhibitors of CDKs are categorized in two families according to their structural similarities and modes of action: the INK4 (named for their ability to inhibit CDK4) family and the Cip (CDK interacting protein)/Kip (kinase inhibitory protein) family.

Recent studies on knockout mice of CDKs and cyclins suggest that one CDK can compensate the other to some extent. For example during the G1 phase the knockout studies revealed that the role of CDK4 and CDK6 could be compensated by CDK2 and in the S phase the role of CDK2 can be compensated by CDK1. However, the appearance of tissue-specific defects in the mutants clearly establishes that they are not totally interchangeable. Moreover, certain tissues are more sensitive to the loss of certain CDKs and cyclins than others (Santamaria et al., 2007).

5.1.5 CDK2

Human CDK2 is a 298 amino acid long enzyme. It belongs to the serine/threonine specific kinases and shares about 50% homology with other CDK family members, including the highly conserved catalytic core region (Hardcastle et al., 2002). The central role of CDK2 in different phases of the cell cycle has already been pointed out above. CDK2-induced phosphorylation of the pRb is the penultimate step before release of the transcription factor E2F and entry into S phase. CDK2 functions throughout the S phase and in G2/M to regulate critical steps in cell growth and division (Wadler, 2001).

CDK2 seems to be the best characterized member of the CDK family in terms of structural biology. There exist approximately 70 X-ray structures of CDK2 and its complexes in the Protein Data Bank (www.rcbs.org/pdb). The high degree of primary sequence homology among members of the CDK family suggests that any one structure of the CDK provides a useful framework for understanding all members of the family.

The structure of monomeric CDK2 has the same overall fold as other eukaryotic protein kinases. The structure consists of an N- terminal lobe rich in β-sheets (five antiparallel β-sheets, β1-5) and a single large α-helix. The larger C-terminal lobe, rich in α-helices contains a pseudo-4-helical bundle (α2, 3, 4, 6), a small ribbon (β6-8) and two additional helices (α5, 7). The crucial ATP binding and catalysis site is found in a deep cleft at the junction of the two lobes (Figure 5.1.3). Importantly, the ATP binding site is present also in the inactive monomeric form. In the monomeric CDK2 structure, two regions are different from the canonical kinase structure and were predicted to function as regulatory elements. One is an α-helix present in other protein kinases but having the unique *PSTAIRE* amino acid sequence only in cyclin dependent kinases; the other is a regulatory loop that has the activating phosphorylation site. This regulatory loop (the T- loop, residues 145-172, containing the Thr160 phosphorylation site) completely blocks the core of the protein substrate-binding cleft. Conserved regions include the glycine rich loop of the ATP binding site (between β1-β2), the highly conserved *PSTAIRE* motif in the α1 helix, the *GDSEID* motif at the N-

terminal end of α5 and the T-loop. The inhibitory phosphorylation sites Thr14 and Tyr15 are both in the middle of the glycine rich loop that serves as a phosphate anchor in ATP binding. The hydroxyl group of Tyr15 is distant from ATP whereas the hydroxyl group of Thr14 is very close to the γ-phosphate of ATP. Thr160 is found at the apex of the T-loop blocking the protein substrate-binding site. The hydroxyl group of the Thr160 is oriented towards the glycine-rich loop of the ATP-binding site (De Bondt et al., 1993).

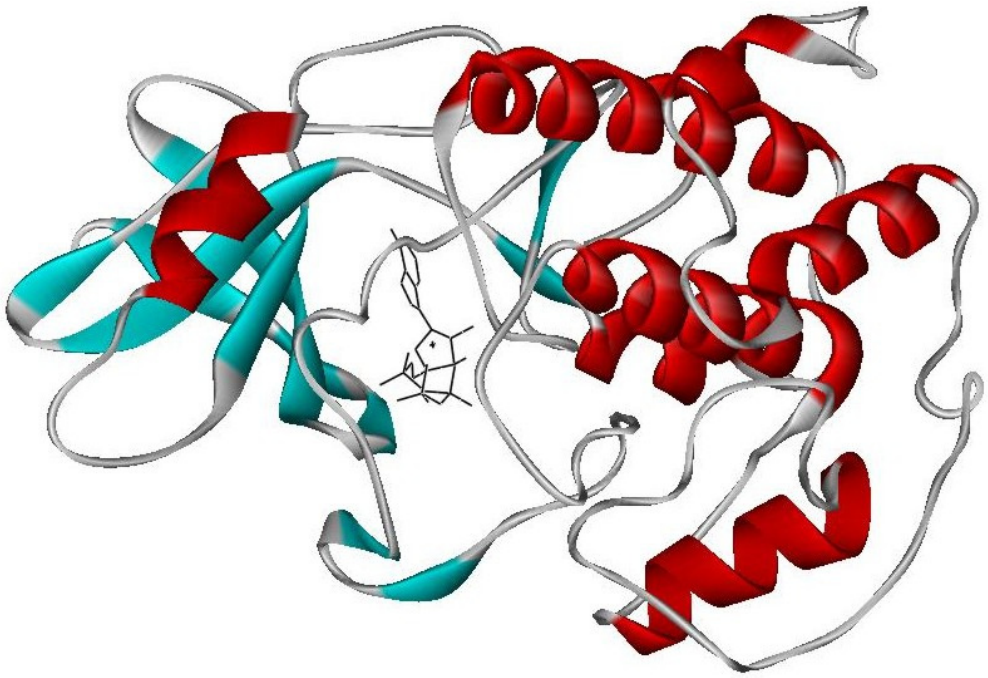

Figure 5.1.3. Overall structure and folding of the CDK2 with ATP bound. The structure consists of the N- terminal lobe rich in β-sheets (shown in blue) plus a single large α-helix (shown in red) and the C- terminal lobe, rich in α-helices. The ATP- (shown as black sticks) binding and catalysis site is found in a deep cleft at the junction of the two lobes (adapted from pdb file 1HCK).

Binding of cyclin A to CDK2 provokes little structural change in the cyclin partner but induces significant modification in CDK2, producing an optimized ATP binding site (Figure 5.1.4). Cyclin A contacts the *PSTAIRE* helix and moves this helix into the catalytic cleft and rotates it by 90°. It also changes the T- loop

structure and position, moving parts of it by over 20 Å. The *PSTAIRE* helix has a catalytic site residue, Glu51, conserved among eukaryotic protein kinases. Without the cyclin the glutamic acid side chain is outside the catalytic cleft. When the cyclin binds, the change in the helix brings the glutamic acid side chain inside the catalytic cleft, where, together with a lysine residue (Lys33), an aspartic acid residue (Asp145), and a magnesium ion it coordinates the ATP phosphate atoms and correctly orients them for catalysis. Asp145, Phe146 and Gly147 form the so called “DFG” motif and are also responsible for inducing the correct conformation in a triphosphate moiety of ATP for catalityc transfer (Hardcastle et al., 2002). Upon cyclin binding the T- loop is moved away and largely relieves the blocking of the catalytic cleft. These changes also expose the phosphorylation site on the T-loop, setting the stage for the full activation of the kinase (Jeffrey et al., 1995).

When the T-loop becomes phosphorylated on Thr160 (CDK2), it undergoes an additional conformational change. This change is induced by the phosphate group acting as an organizing center in this region being bound by 3 arginine side chains, each coming from a differenet part of the structure (one from the N- lobe, one from the C- lobe, and one from the T- loop). The arginine residues are in turn hydrogen bond to other CDK and cyclin groups and extend the organizing influence of the phosphate group. The significance of this CDK region is that it is part of the catalytic cleft, and in particular the putative polypeptide substrate interaction site. Phosphorylation thus completes the re-organization of the substrate-binding site that was started by the cyclin (Jeffrey et al., 1995; Russo et al., 1996).

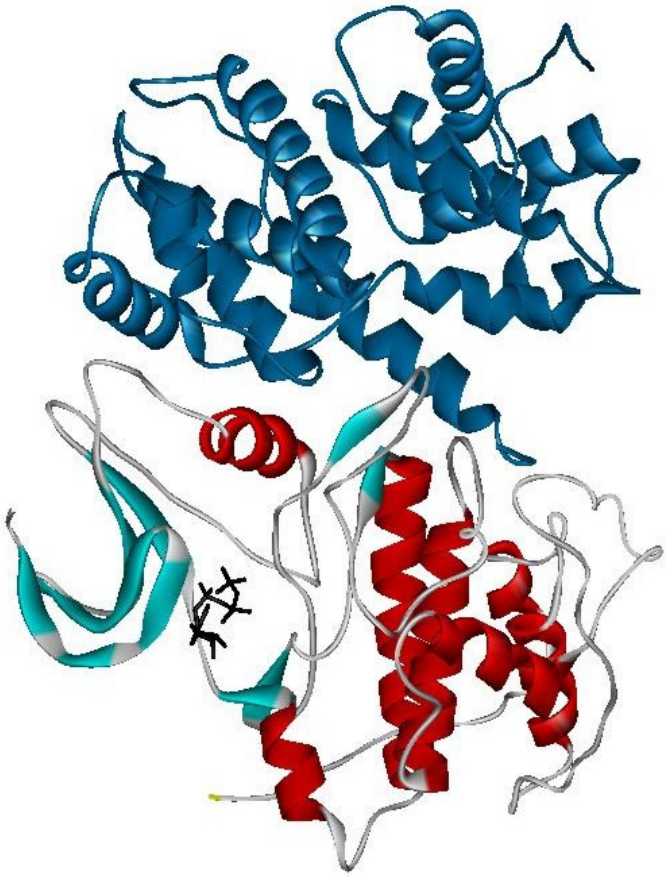

Figure 5.1.4. Overall structure and folding of the binary complex of cyclin A-CDK2 with ATP. Cyclin A shown in dark blue, CDK2 as on Figure 5.1.3.

5.1.6 Inhibitors of CDK2

As already mentioned one of the important ways of regulation of CDK activity is specific inhibition. The main group of protein inhibitors for CDK2 namely p21, p27 and p57 belongs to the Cip/Kip family, categorized according to their structural similarities and modes of action. These proteins inactivate the CDK-cyclin complex. The other endogenous proteins and their actions which influence the CDK2 activity apart from the Cip/Kip family are described in Table 5.1.2 (Meijer et al., 2000).

Table 5.1.2. Endogenous inhibitors of CDK2 and their roles in the cell cycle

Inhibitor	Action
MYC	Oncoprootein that induces CDK2 activity. Down-regulation of MYC induces cell cycle arrest by down-regulation of CDK2
Interferon/TNF (Tumor Necrosis Factor)	Reduces CDK2 activity by the combination of cytokines
Insulin receptor kinase	CDK2 activity is inhibited as this protein recruits p27
PKC-protein kinase C	Association with cyclinA-CDK2 ultimately leads to the inhibition of kinase activity

Frequent deregulation of CDK2 in cancers has lead to the search for different compounds that can specifically inhibit this kinase. 6-dimethylaminopurine and isopentyladenine were the first non- physiological inhibitors of CDKs. The potential of pharmacological CDK2 inhibition is already widely explored. The selectivity of these compounds decreased in clinical trials. This is because the amino acids constituting the ATP binding region in the CDK family are more or less conserved while all proposed CDK2 inhibitors are small, hydrophobic and can bind competitively to the ATP binding site of CDKs.

Until today a large number of different inhibitors of CDK2 are known. Many of them have been co-crystallized with CDK2 (Meijer et al., 2000). Some of these inhibitors are being tested in clinical trials, for example flavopiridol, roscovitine or UCN-01 (Achenbach et al., 2000; Senderowicz, 1999; Senderowicz, 2003). In pre-clinical and clinical studies, flavopiridols induced programmed cell death, promoted cell differentiation, inhibited angiogenesis, and modulated transcriptional events (Motwani et al., 2000; Robey et al., 2001). Extensive clinical trials were carried out with the purine inhibitor, roscovitine which in combination with chemotherapy and radiotherapy can be helpful in arresting the cell cycle. UCN-01 (7- hydroxystaurosporine) inhibits CDK2 in a concentration dependent fashion. UCN-01 acts via p21 (Wang et al., 1996); it induces the expression of p21, which finally stops the kinetic activity of CDK2. UCN-01 has

been successful in clinical trials (Sausville et al., 2001; Bunch et al., 1996; Akiyama et al., 1997). Another group of substances that have been reported to show promising therapeutic efficacy against human cancers are chalcones, derivatives of flavonoids and isoflavonoids (Stoll et. al., 2001; Kumar et al., 2003). To this group belongs also butein (3,4,2',4'-tetrahydroxychalcone), a plant phenol which is one of the major biologically active components found in the bark and stems of *Rhus verniciflua* (Yang et al., 1998).

5.1.7 Cell cycle and cancer

As cancer is often regarded as a disease of uncontrolled cell growth many of the cell cycle regulatory mechanisms have been found to be unsettled in tumors. Cell cycle deregulation associated with cancer occurs through mutation of proteins important at different levels of the cell cycle. It can be either mutations in so called proto-oncogenes or in tumor suppressor genes. Mutated versions of proto-oncogenes which in normal cells are responsible for stimulation of the cell proliferation can promote tumor growth. Inactivation of tumor suppressor genes like pRb and p53 results in dysfunction of proteins that normally inhibit cell cycle progression. In cancer, mutations have been observed in genes encoding CDK, cyclins, CDK-activating enzymes, CKI, CDK substrates, and checkpoint proteins (Sherr, 1996).

The p53 gene is the most frequent site of mutation found in human cancers. p53 is involved in a DNA-damage cell-cycle checkpoint pathway and is able to induce either cell cycle arrest or apoptosis (Jayaraman et al., 1999). Point and missense mutations lead to conformational changes and inactivation of the protein. The p16 inhibitor protein and tumor suppressor has been found to be mutated in approximately one third of human cancers. Cells with altered p16 will be unrestrained to proceed through G1 (Vermeulen et al., 2003). A few modifications have been found in p18 and p21 in breast tumor and thyroid carcinoma respectively (Lapointe et al., 1996; Shi et al., 1996). Low expression levels of the p27 inhibitor protein appear in many tumors and correlate with poor prognosis while overexpression of cyclin D is commonly observed in breast cancer (Hardcastle et al., 2002). Mutations in CDK4 and CDK6 genes, resulting

in loss of CKI binding, have also been identified (Easton et al., 1998). CDK1 and CDK2 have been reported to be overexpressed in a subset of colon adenomas and a greater overexpresion was seen in focal carcinomas in adenomatous tissues (Kim et al., 1999). Mutation of CDK4 that occurs as a result of amplification has been observed in melanomas, sarcomas, and gliomas (Woelfel et al., 1995). Recently the *CDK8* gene was found in a frequently amplified region of a human genome in colon cancer, suggesting that its increased level of expression contributes to the development of colorectal tumors (Firestein et al., 2008). The latest news is that the CDK8 orchestrates cross-talk between two signalling pathways that are frequently deregulated in human cancers. As well as the direct stimulatory effect of CDK8 on β-catenin's transcriptional activity, CDK8 seems to have a second (less direct) stimulatory effect on β-catenin, by protecting it from inhibition by E2F1 (Bernards, 2008).

The evidence that CDKs, their regulators and substrates are disordered in many human cancers has stimulated the search for chemical CDK inhibitors as potential cancer therapeutic. Different strategies of therapeutic intervention were investigated: targeting the major regulators of CDK activity including overexpression of CKI, synthesis of peptides mimicking the effects of CKI, decrease of cyclin levels, modulation of the phosphorylated state of CDK and of the enzymes regulating it (McDonald and elDeiry, 2000). But until now the direct inhibition of CDK kinase activity by competitive inhibition of ATP binding has been the most successful strategy for the development of a potential therapy (Vermeulen et al., 2003). The CDK inhibitors help to trigger apoptosis in tumor tissue. CDK inhibitors also protect normal cells from apoptosis due to the toxic effects of chemotherapy. Some inhibitors were found to prevent chemotherapy-induced alopecia by arresting the cell cycle and reducing the sensitivity of the epithelium to many active oncogenic agents (Davis et al., 2001). CDKs, which have been explored, are CDK1, CDK2, CDK4/6. Recent studies are focusing on the other kinases like CDK7, 8, 9 and the selectivity of the potential inhibitor is a key issue. So far no inhibitor has been developed that will inhibit the cell cycle by specifically inhibiting a particular CDK.

5.2 Goals of the study

The CDKs play a key role in controlling the cell cycle and apoptosis. Malfunction of the cell cycle and its regulatory mechanisms, especially CDK activity, is frequently found in many human tumor cells. Highly selective inhibitors of CDK are of great potential to be cancer therapeutics. Structural analysis of CDKs and their inhibitors improves our understanding of the structure-function relationship of this multidimensional protein family. A precise model of the interaction between CDKs and their binding partners can be created based on high-resolution 3D structures of the interacting proteins, kinetic data from binding studies on truncated variants, and information on their dynamic behavior from NMR experiments.

The first goal of the work presented in this part of the thesis was to study the molecular basis of action of a new small molecule inhibitor for CDK2 by means of x-ray crystallography. As CDK2 and its action is already well examined as well as a number of its potential binding partners are known, studying CDK2 – inhibitor interaction was chosen to illustrate a novel NMR-based drug screening method. The challenge was to permit rapid screening by monitoring a simple 1D proton NMR signal of the NH side chain ($^{N}H^{\varepsilon}$) of the tryptophans. The method also provides quantitative characterization of the antagonist-protein and antagonist-protein-protein interactions in the form of K_Ds and fractions of the released proteins from their mutual binding.

5.3 Results

This section of the thesis contains results of work on CDK2 and its binding partners. It involves the crystallization trials as well as the illustration of the new approach to NMR investigation of protein-inhibitor interactions.

5.3.1 Construct design and expression of CDK2

All constructs of CDK2 were based on previously published literature (Bondt et al., 1993; Majumdar, 2005). Various mutants of CDK2 were designed and cloned in order to perform NMR experiments (Rothweiler, 2006). Several tryptophan residues were introduced in the positions next to the binding site of ATP so that they were sensitive to the binding of tested inhibitors (Figure 5.3.1). The list of the constructs of CDK2 used in this work is shown in Table 5.3.1.

Table 5.3.1. List of CDK2 constructs used in this work

	Construct name	Vector	Expression in *E. coli*
1	**CDK2pro**	pProExHTa	high, soluble
2	**CDK2wild**	pET-28a	high, soluble
3	**CDK2S53W**	pET-28a	high, soluble
4	**CDK2H84W**	pET-28a	high, soluble
5	**CDK2A93W**	pET-28a	high, soluble

Properties of the designed constructs, such as their theoretical isoelectric points (pI), molecular masses (MW) and extinction coefficients (E_{280}) are shown in Table 5.3.2. This information was deduced on the basis of the amino acid composition with the aid of the program ProtParam on www.expasy.org (Gasteiger et al., 2003). Examination of established methods of purification of similar or homologous proteins was used in designing purification protocols.

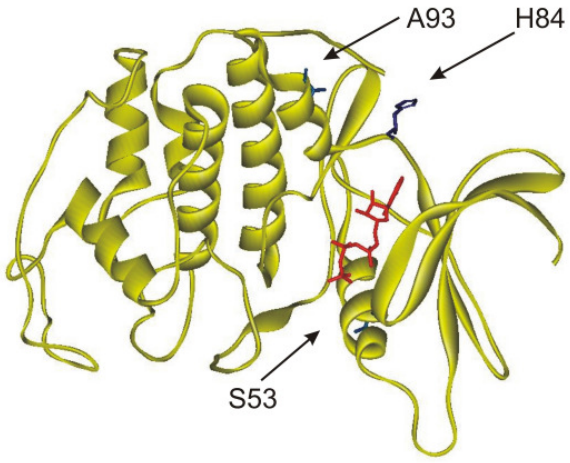

Figure 5.3.1. Structural organization of CDK2. Point mutations introduced for NMR experiments are shown in blue; the bound ATP is shown in red (Jeffrey et al., 1995)

Table 5.3.2. Physicochemical properties of the investigated constructs

	Construct name	N. of aa	MW (Da)	E_{280} ($M^{-1}cm^{-1}$)	pI
1	**CDK2pro**	323	37141	48485	7.79
2	**CDK2wt**	331	37386	37025	9.08
3	**CDK2S53W**	331	37485	42525	9.08
4	**CDK2H84W**	331	37435	42525	9.08
5	**CDK2A93W**	331	37501	42525	9.08

5.3.2 Purification of CDK2

Purification of all CDK2 constructs followed a similar protocol, involving two different liquid chromatography steps: immobilized metal affinity chromatography on Ni-NTA in the native conditions followed by gel filtration chromatography.

5.3.2.1 Affinity chromatography (Ni-NTA)

All expressed proteins were 6-His-tagged for nickel affinity chromatography under denaturing and reducing conditions. The pH of the supernatant was adjusted to 8.0 with 1 M NaOH, which is optimal for the binding

to the Ni-NTA resin. A Ni-NTA slurry was added and binding was performed at 4℃ for 2 h with gentle agitation.

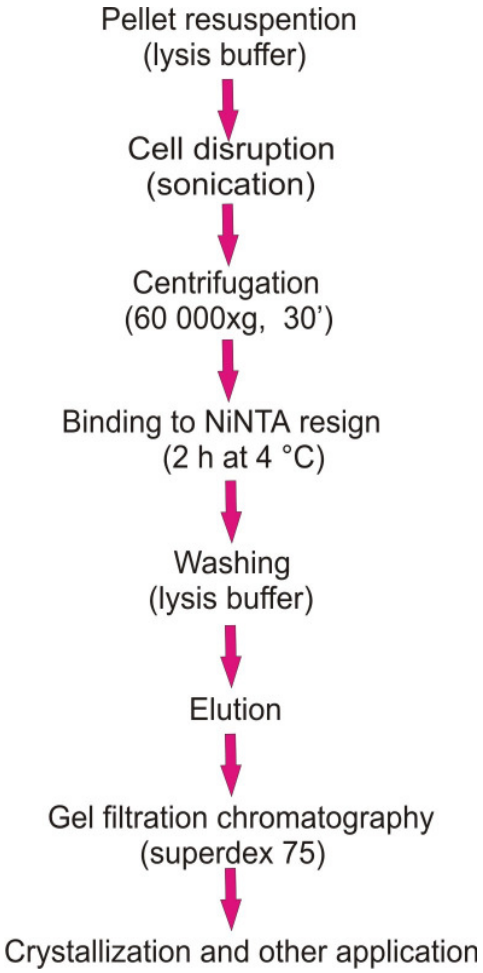

Figure 5.3.2. Flow chart of the purification scheme for all of the CDK2 constructs.

5.3.2.2 Gel filtration chromatography

Gel filtration was the final step of purification to separate the protein from possible aggregates. Complete exchange of the buffer was also possible, removing any low molecular weight substances that could interfere with NMR measurements. The 120-ml Superdex 75 prep grade column was chosen due to its good resolution in the range 10-75 kDa. The samples loaded onto the column were not larger than 6 ml. A typical chromatogram and the SDS-PAGE of the CDK2 purification are shown in Figure 5.3.3.

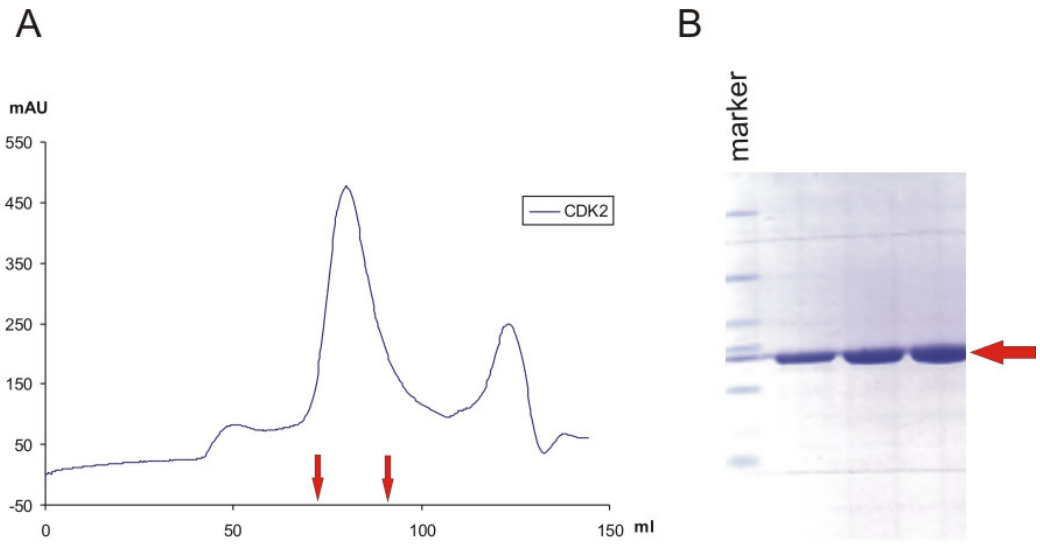

Figure 5.3.3. A typical chromatogram and the SDS-PAGE analysis of CDK2 after gel filtration. Pooled fractions as well as purified protein are indicated by red arrows.

5.3.3 Functional and structural studies

5.3.3.1 NMR studies of the folding and organization of CDK2 constructs

All investigated constructs were examined by NMR. 1D spectra were taken. Inspection of such spectra yields semi-quantitative information on the extent of the folding in partially structured proteins or their domains (Rehm et al, 2002). The large signal dispersion beyond 8.3 ppm proves a protein to be folded. The appearance of intensities at chemical shifts near ~8.3 ppm is an indicator for a disordered protein, as this is a region characteristic of backbone amides in a random coil configuration (Wüthrich, 1986). The one-dimensional proton NMR spectra of various mutants of CDK2 are shown in Figure 5.3.4. The spectra show that all mutants are folded in solution.

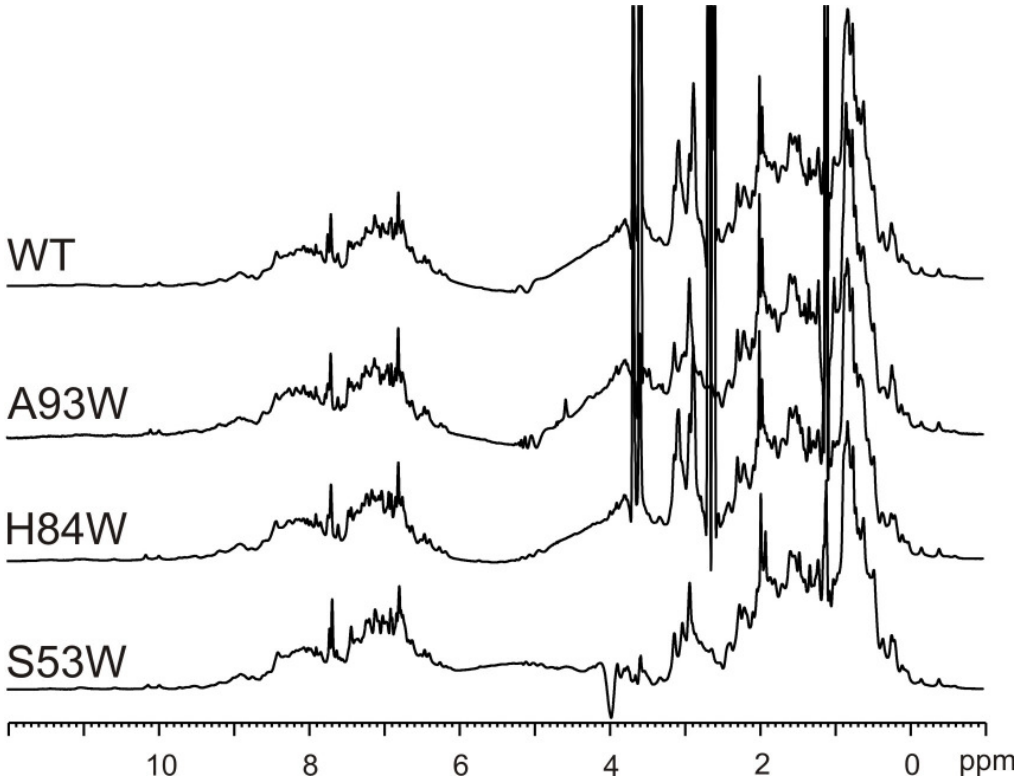

Figure 5.3.4. Characterization of the structural integrity of the CDK2 wild type and its various mutants by one-dimensional proton NMR spectroscopy. Very sharp signals are due to buffer impurities.

Close inspection of NMR spectra at a chemical shift of around 10 ppm show also that the indole side-chain of the introduced Trp is flexible, with a linewidth ($\Delta\nu_{exp}$) of 19±2 Hz, whereas the $\Delta\nu_{exp}$ of the NH indole of native Trp´s is 29±4 Hz and the estimated $\Delta\nu_{exp}$ of backbone amides between 50-60 Hz (Figure 5.3.4; Figures 5.3.8-10).

While for the S53W mutant the signal was broadened and weak, the H84W and the A93W mutant gave sharp signals. The newly introduced tryptophan of the H84W mutant overlaps with one endogenous tryptophan whereas the signal of the tryptophan of the A93W mutant was well separated from the rest of the $^{N}H^{\varepsilon}$ Trp signals.

5.3.3.2 Comparison of the folding and stability of CDK2 and its mutants by thermal shift assay

The stability of a protein in solution can be dramatically influenced by introduced mutations. This phenomenon causes changes in protein melting temperature. By utilizing special fluorescent dye, quenched in a hydrophobic environment but active in the proximity of exposed hydrophobic core of the protein, the thermal unfolding process can be monitored. Thus, investigation of stability of a protein and its mutant can be performed by comparison of the melting temperature of both. Obtained results indicated no relevant changes in stability of particular mutants. Their melting points were similar at around 40℃ (Figure 5.3.5).

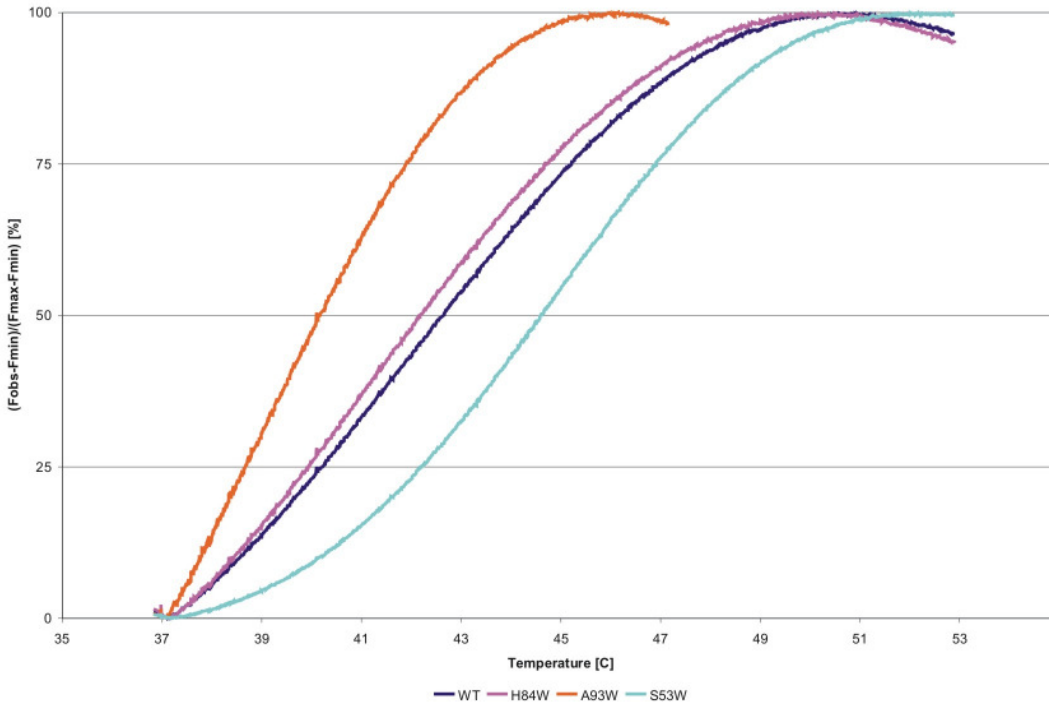

Figure 5.3.5. Fluorescence graphs showing melting characteristics of wt-CDK2 and its mutants. The melting points are at around 40℃.

5.3.3.3 CDK2 binding partners

A number of CDK inhibitors of considerably different chemical classes have been reported (Huwe et al., 2003). In order to the performe NMR and crystallography binding studies, the following CDK2 inhibitors were used: roscovitine, butein and compound 19 (Figure 5.3.6)

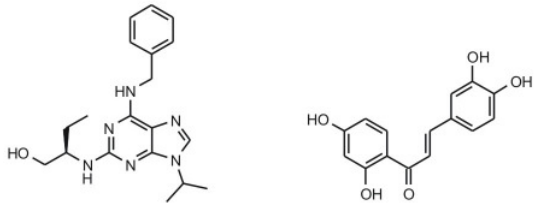

Figure 5.3.6. Inhibitors for CDK2.

Roscovitine (seliciclib, CYC202) is a strong binding CDK2 inhibitor with an IC_{50} of 0.1 µM. It is one of the most extensively tested inhibitors (Meijer et al., 1997 and 2003; Whittaker et al., 2004). Roscovitine is an orally bioavailable compound with high selectivity for CDK1, 2, 7 and 9 (Shapiro, 2006). Preclinical studies have shown antitumor effects in a broad number of human tumor xenografts and phase I clinical trials are ongoing (McClue et al., 2002; Raynaud et al., 2005; Benson et al., 2007).

Butein was chosen as an example of a weak inhibitor for CDK2. Butein (3,4,2',4'-tetrahydroxychalocne), a plant polyphenol, was first described as a novel cAMP-specific phosphodiesterase inhibitor, which produced endothelium dependent relaxation of rat aorta (Yu et al., 1995). It inhibits the tyrosine specific protein kinases, like the EGF receptor tyrosine kinase and $p60^{c\text{-}src}$ with an IC_{50} of 16 µM and 65 µM, respectively, but has little effect on serine-threonine specific kinases, PKC and PKA (IC_{50} >500 µM) (Yang et al., 1998).

Compound 19 for crystallization trials was received from a pharmaceutical company, its identity was not disclosed.

5.3.3.4 The new approach to NMR screening for CDK2 inhibitor binding

The general idea of using NMR spectroscopy to detect protein-ligand interaction is the observation of a changed chemical shift of an identified residue of the protein due to a novel chemical environment provided by the presence of the ligand. 2D NMR methods are often used for binding studies although they are time consuming and expensive (Shuker et al., 1996) Due to the large number of signals it is generally impossible to identify the peaks of a specific residue in a 1D spectrum, which would be much faster and cheaper. The $^{N}H^{\varepsilon}$ signal of Trp is an exception to this rule as its chemical shift at around 10 ppm lies far away from the signals of the other amino acids. Therefore observation of Trp signals, either endogenous or introduced via protein engineering provides insights in ligand binding by simple 1D NMR spectroscopy (see also Chapter 5.4 for details).

The full-length wild type CDK2 protein has four endogenous tryptophans within its amino acid sequence. Since CDK2 is relatively large, the tryptophans, which are all buried inside the molecule, give rise to broadened NMR signals and, although well separated from each other, they were not sensitive to the binding of inhibitors that have been tested. Other regions of the CDK2 spectrum could also not provide unambiguous conclusions on ligand binding (Figure 5.3.7c).

To employ the new Trp-NMR strategy, a number of CDK2 mutants were designed to introduce a Trp residue close to the binding site (Figure 5.3.1). First the A93W mutant was titrated with roscovitine. A stepwise addition of roscovitine leads to a splitting of the new tryptophan peak, indicating strong binding (Figure 5.3.7a).

Next butein, a weak inhibitor for CDK2, was tested. A stepwise addition of butein to the A93W CDK2 mutant lead to a continuous movement of the newly introduced tryptophan signal (Figure 5.3.7b). This experiment shows that butein is a weak inhibitor for CDK2 (K_D: 100±50 µM).

Identical behavior was seen in titrating the H84W mutant with roscovitine (Figure 5.3.8). Due to the overlap with the endogenous peak at around 10.15 ppm, the starting peak is a sum of 84W and an endogenous tryptophan.

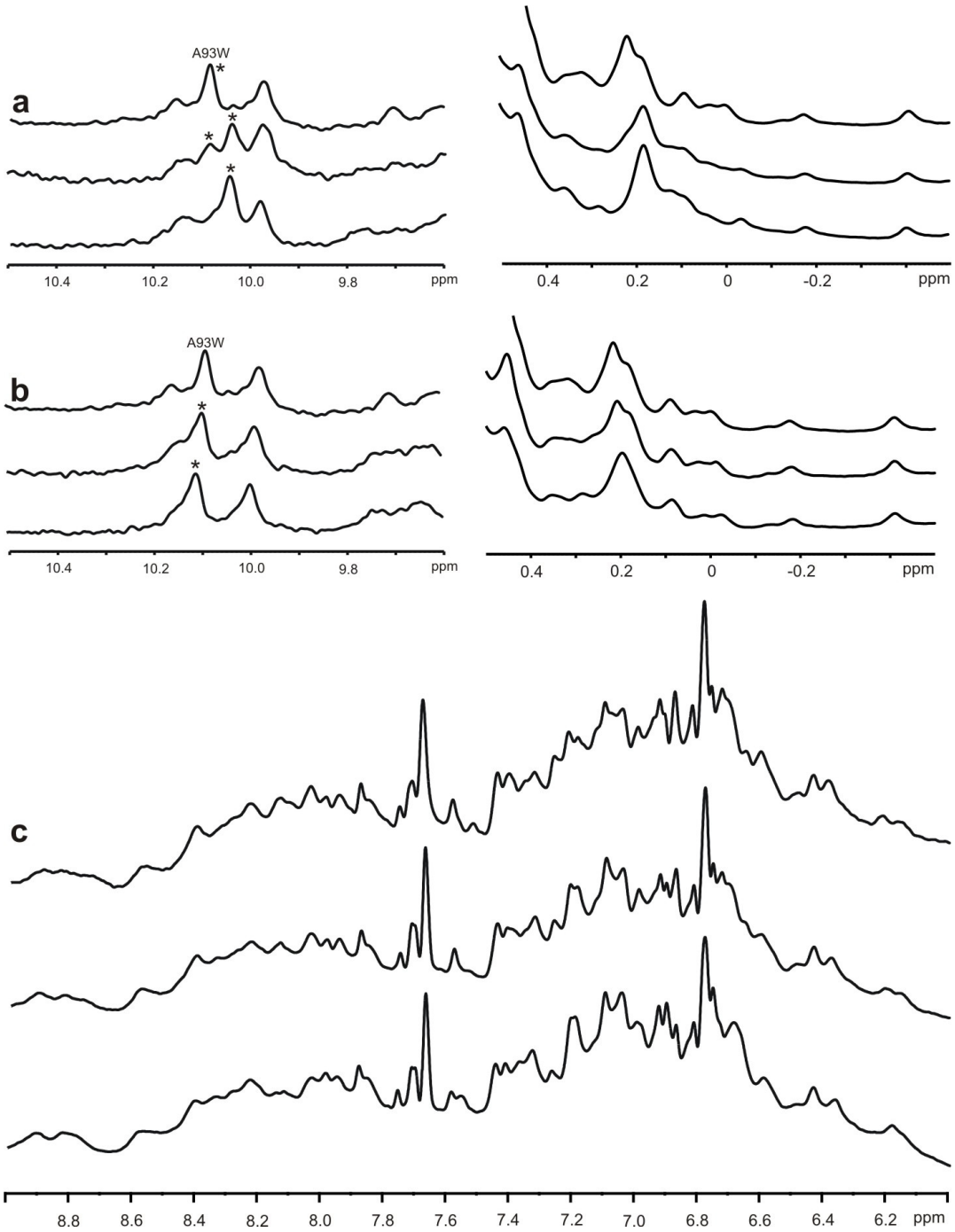

Figure 5.3.7. NMR spectra of the CDK2 mutant A93W titrated with increasing concentrations of roscovitine (a) or butein (b). The left side shows the region with the new introduced tryptophan. The right side shows the aliphatic region. Whereas the NMR signal of the introduced tryptophan shows clearly the differences between strong and weak inhibitors, the peaks in the aliphatic regions are difficult to interpret and differentiation between strong and weak inhibitors is

Figure 5.3.7 (cont.)

not possible. Top spectra: Unbound CDK2; middle: molar ratio protein: inhibitor 1: 5; bottom: molar ratio protein: inhibitor 1: 10. The asterisk indicates the new introduced tryptophan 93W. (c) The amide region of the CDK2 mutant A93W titrated with roscovitine. The amide region is so crowded that it is impossible to trace individual signals and differentiate between strong and weak inhibitors.

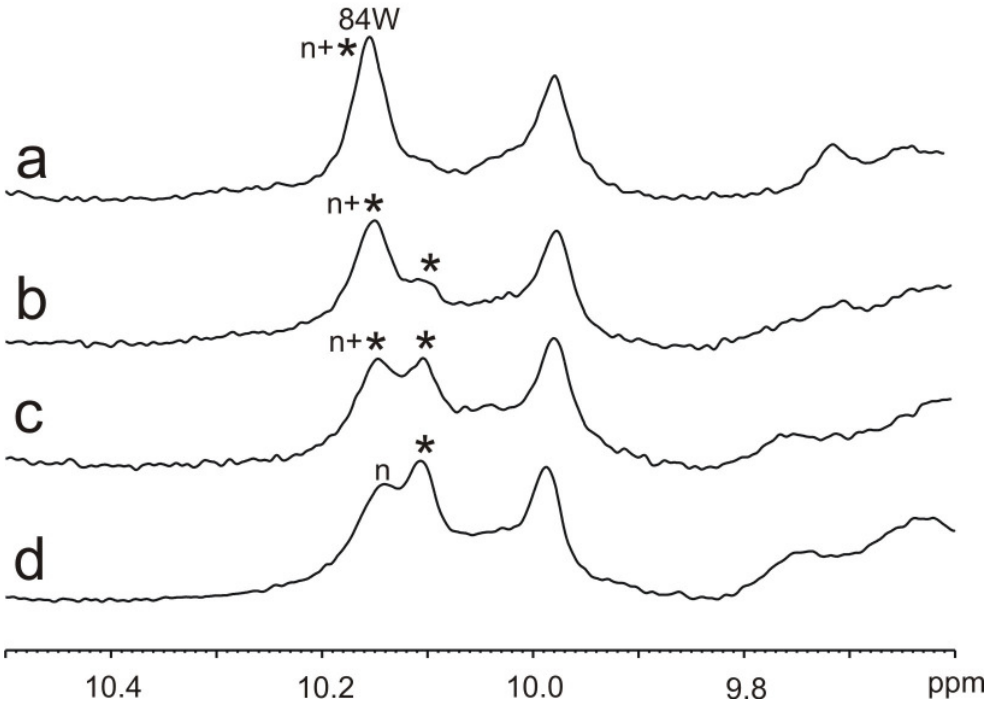

Figure 5.3.8. 1D proton NMR spectra of human CDK2-H84W titrated with roscovitine. (a) Free CDK2. (b) The molar ratio protein: inhibitor, 1: 0.3. (c) Molar ratio protein: inhibitor, 1: 0.6. (d) Molar ratio protein: inhibitor 1: 1. The asterisk indicates the new introduced tryptophan 84W, the endogenous tryptophan is marked with a small n.

5.3.3.5 Crystallization trials of CDK2 in complex with compound 19

5.3.3.5.1 Optimization of crystallization conditions

The protein prepared for crystallization was separated from any possible degradation products and low molecular weight substances by gel filtration chromatography. Gel filtration was also used for a thorough buffer exchange. The solvent used contained low salt (15 mM NaCl) and buffering substance (10 mM Hepes, pH 7.4) to minimize any possible interference with crystallization reagents. Purity, homogeneity and folding of the proteins to be used in crystallization trials were evaluated using mass spectrometry, amino-terminal sequencing and NMR.

The crystallization of CDK2 was successful with 50 mM ammonium acetate, 0.1 M Hepes, pH 7.4 and 15% PEG 3350. Crystals appeared within 2 weeks (Figure 5.3.9), in drops containing relatively high concentrated protein (~10 mg/ml). In the drops in which the concentration of the protein was lower (5-7 mg/ml) only precipitate was observed. Condition for growing crystals of CDK2 was further optimized in order to achieve crystals of better quality. For this purpose, crystallization was carried out with buffers in which the concentration of PEG 3350 and pH of Hepes were varied (10%-19% for PEG 3350 and 7.0-7.8 for Hepes). Finally the concentration of CDK2 used for crystallization was about 8 mg/ml. Additionally, the micro-seeding technique was employed. The crystal seeds were introduced into equilibrated clear drops using a rabbit's whisker. The new crystals grew along the streak line within two days.

In the initial experiments, native crystals of CDK2 were soaked by direct addition of inhibitor solution to the crystallization wells. Later the CDK2-compound 19 complex was prepared before crystallization by mixing equimolar amounts of the components but, in this case crystallization was not successful.

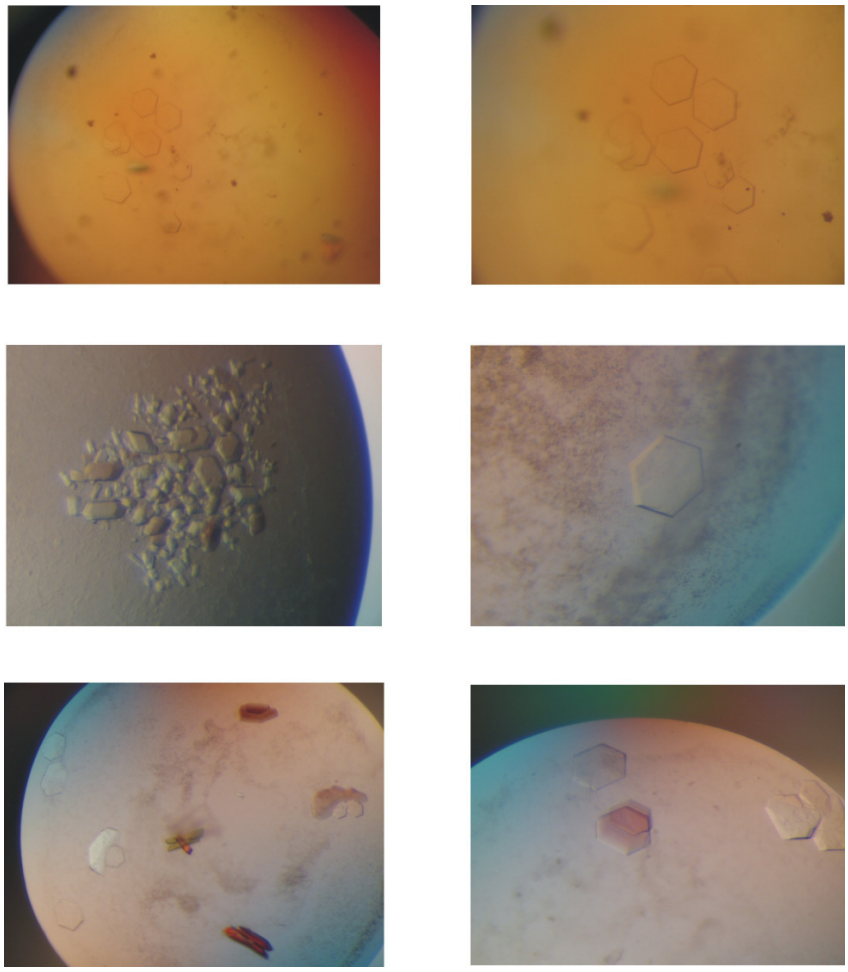

Figure 5.3.9. Crystals of CDK2 grown in 50 mM amonium acetate, 0.1 M Hepes, pH 7.4 and 15% PEG 3350 at 4℃.

5.3.3.5.2 Data collection

X-ray data sets were collected at the MPG/GBF beamline BW6 at DESY, Hamburg**.** For protection, all measured crystals were soaked for 1-2 min in a solution containing 80% (v/v) mother liquor and 20% (v/v) ethylene glycol and frozen in liquid nitrogen. The measurements yielded high quality data sets, the best one up to 2.1 Å resolution.

Data processing was done with MOSFLM (Leslie, 1991) or the XDS software package (Kabsch, 1993). The structure of the crystal was determined by molecular replacement using MolRep from the CCP4 program suit. As search model CDK2 (PDB ID: 2uzo) was chosen. The space group was $P2_12_12_1$ with unit cell dimensions a = 53.5 Å, b = 71.4 Å, c = 72.3 Å, α = β = γ = 90°, The model

was refined with good agreement to standard stereochemical parameters. Unfortunately, further structural analysis has not shown the bound compound 19.

5.4 Discussion

Human CDK2 is a 34 kDa serine/threonine specific kinase (Bondt et al., 1993). It plays a pivotal role in different phases of the cell cycle and its deregulation contributes to tumorigenesis. The potential of CDK2 targeting by small molecule compounds as a cancer therapy is already widely explored. There exist approximately 70 X-ray structures of CDK2 and its complexes in the Protein Data Bank (www.rcbs.org/pdb) and a large number of different inhibitors of CDK2 are known. Some of these inhibitors, like for example flavopiridol and roscovitine, are being tested in clinical trials (Achenbach et al., 2000; Senderowicz, 2003). One of the aims of the study presented in this section was to crystallize CDK2 with a new potential inhibitor, called “compound 19”, provided by a pharmaceutical company. The crystallization of CDK2 was successful with 50 mM ammonium acetate, 0.1 M Hepes, pH 7.4 and 15% PEG 3350. Unfortunately, in the initial experiments where native crystals of CDK2 were soaked by direct addition of the inhibitor solution to the crystallization wells, the structural analysis didn’t show the compound.

It is a common experience that even when conditions are well established for the growth of crystals of a protein of interest, obtaining co-structures with new inhibitors can be difficult (Danley, 2006). Crystal soaking is a technique widely used in the study of enzyme active site inhibitors, however, too often it is not successful. When compared with small-molecule crystals, protein crystals are loosely packed, typically containing from 30 to 80% solvent (Matthews, 1968; Matthews, 1974). The network of lattice interactions determines the size and configuration of channels traversing the crystal, which average typically from 20 to 100 Å in diameter (Vilenchik et al., 1998). These channels contain the bulk solvent bathing the crystals in addition to a shell of `bound water' interacting with the protein molecules and provide considerable access for small ligands to the protein molecules in the lattice. The problem of soaking experiments might be the not optimal accessibility of the binding pocket on the crystallized protein, though

in this case the binding affinity of the inhibitor was most probably not strong enough to penetrate the highly structured crystal lattice of CDK2.

NMR-based drug screening methods, in particular the NMR chemical shift perturbation method, are powerful techniques used for the identification of ligand-protein interactions and provide the most reliable characterization of binding propensities of ligands to their target proteins. (Jahnke et al., 2004; Lepre et al., 2004; Peng et al., 2004). They are, however, one of the least effective methods in terms of the amount of protein required and the time needed for running an NMR experiment (Shuker et al., 1996). The now widely used screening approach "SAR by NMR" employs typically ^{15}N-labeled proteins that are monitored by 2D ^{15}N-^{1}H-correlation NMR spectroscopy (Shuker et al., 1996; Stockman and Dalvit, 2002). Approximately 0.5 ml of a 0.2 mM protein solution (Meyer and Peters, 2003) is required for a 15 kDa protein per tested ligand or mixture of ligands. Lower concentrations of protein can be used, but this increases the expense of the time needed to record a qualitatively acceptable spectrum. The microcoil-probe NMR technology, enhanced with the cryoprobe detection, can help in lowering the amount of protein needed to microgram quantities (Aramini et al., 2007), but still the performance of NMR is slow compared to other high-throughput techniques, like for example, fluorescence spectroscopy (Buehler et al., 2005).

The work presented in this section of the thesis is part of a research project for a novel NMR-based drug screening method. Complete results were recently published (Rothweiler et al., 2008). The method is illustrated on inhibitors of CDK2 and the p53-Mdm2 interaction (Vousden and Lane, 2007; Chene, 2003; Vassilev, 2006). The results presented here shows that simple 1D proton NMR spectra of unlabeled proteins (i.e. no ^{15}N or ^{13}C labeling required) can be utilized in screening for lead compounds thus permitting rapid ligand characterization. Since signal overlap in proton 1D spectra of proteins may present a problem, one must have or must introduce an amino acid "reporter" that has at least one not overlapped NMR signal which is sensitive to the binding of the ligand to the investigated protein. Here it has been chosen to introduce

tryptophan, because it is the only amino acid whose NH indole side chain ($^{N}H^{\varepsilon}$) gives an NMR signal at around 10 ppm at physiological pH. The signal is hence well separated from the bulk of the amide protons and can be easily monitored (Wüthrich, 1986).

In principle, one can use any 1D resolved signal of proteins, including backbone amides and methyl groups, to detect ligand binding. In most practical cases, however, it is not possible to use amide and methyl signals for proteins larger than ca. 10 kDa, because these signals are not resolved in 1D proton NMR spectra and/or might not be sensitive to ligand addition. CDK2, although relatively large (298 amino acids), gave exceptionally clear spectra in the methyl region (Figure 5.2.4). However, also here the shifts upon ligand binding of NMR peaks are so complicated that no unambiguous conclusion can be reached from the data. Small molecule binding is usually in a µM range and a large excess of these compounds must be used to observe the binding to the protein (binary interactions) or inhibition of the protein-protein interaction (NMR competition experiments on ligand/protein complexes). The signals of a compound obscure usually the signals from the protein. This gets worse, if mixtures of compounds are used. In addition, many potential compounds require 0.3% of detergents for solubilization. These types of detergents like Tween 20 or CHAPS totally obscured aliphatic NMR signals.

Presented experiments show that, by introducing tryptophan residues in surface exposed sites near potential antagonist binding sites, small molecule binding can be monitored through effects on the indole proton chemical shifts. Three residues on CDK2 have been mutated. These mutations change only insignificantly protein stability, fold, and binding capabilities. NMR spectra clearly show that side-chains of tryptophans in the CDK2 mutants are flexible and get restricted in motions upon binding of a ligand (Figure 5.3.1). For example, the tryptophan in the A93W mutant is part of a helix, which is situated parallel to the binding pocket for the ligands and should thus have its side chain pointing outward without steric clashes with the rest of the protein. After close inspection of the models of the structures of the CDK2 mutants, it is not expexted that the

binding site is altered by the mutation. Therefore it should bind to ligands in an approximately same way as wt-CDK2.

Two mutants of CDK2, A93W and H84W, proved to be universal and excellent reporters of all interactions of small molecules which were studied. The A93W mutant gave the best results in distinguishing between strong and weak binders. The H84W mutant gave similar results, the overlapping of the introduced tryptophan signal with one of the four naturally present tryptophan signals required, however, more careful analysis of the binding data. The detection of strong and weak binders by introduction of tryptophans to the CDK2 sequence, a protein which already has four tryptophans in the sequence, shows that the method could be expanded to larger proteins, which are expected to have many natural tryptophans. Even coincidental overlap with a naturally present tryptophan signal is not detrimental.

In some cases the naturally present tryptophans might be used for detecting the binding without the need of introducing extra tryptophans. This is the case when tryptophans actively participate in the binding to ligands or are close enough to the binding pocket and are therefore sensitive to NMR experiments. This is illustrated by the titration of the protein Mad2 with Mad1 (see section 4.3).

The method described here can be used for many proteins. The design of point mutations needs structural information for a protein of interest, which is however in most cases available. The presented method can easily be combined with 2D NMR methods that use labeled proteins. For example, ^{15}N labeled Trp mutants can be used for the 1D Method, if one puts up with the small inconvenience of the NH splitting of the Trp indole signals. Of course in 1D spectroscopy for the sake of speed, one has to sacrifice spectral simplicity. Thus, mapping binding interfaces in proteins is best achieved with 2D NMR and $^{15}N(^{13}C)$ labeled proteins. This can be achieved with the “traditional” SAR by NMR (Shuker et al., 1996; Klages et al., 2006) or using a recently devised, elegant assay, IDIS-NMR, which allows for simultaneous 2D NMR monitoring of

two protein components of the complex in a single sample (Golovanov et al., 2007).

2D NMR-based 'SAR-by-NMR' approaches that employ ^{15}N-labeled proteins take typically several hours per one spectrum. With the presented method this time can be cut to few minutes. For example, it took 8 min to obtain a 1D spectrum with 128 scans with reasonable S/N of the CDK2 mutants. A four-step titration can be recorded in 30 min. For comparison, an average ^{1}H-^{15}N HSQC spectrum has taken 5 h (128 scans, 96 increments) for CDK2 samples of the same concentration as that used in the 1D experiment. The developed technology can potentially reform the process of NMR-based screening for lead compounds allowing for high-throughput screening pipelines. It should also allow analysis of proteins from low-yield protein production systems, such as cell-free, insect cell and other eukaryotic expression hosts. Importantly, the method can also provide a rapid and straightforward quantitative characterization of the antagonist-protein in the form of K_Ds and fractions of the released proteins from their mutual bindings. The features of the technologies described here should thus have useful applications in lead generation in drug discovery.

6 Summary

This thesis is a collection of three different projects to characterize the biochemical and structural features of proteins important in cancer development. The wide range of open questions in cancer research is reflected in the diversity of the projects.

A combination of X-ray crystallography and NMR spectroscopy could provide structural insights to establish a model of the IGF-1R activation by IGF. To arrive at a crystall structure extensive trials with various fragments of different length and mutations of IGF-1R were undertaken. All experiments, including the performed expression and purification steps as well as initial NMR examinations are described in detail to aid future attempts.

The elucidation of the three dimensional structure of the complex of Shugoshin and Mad2 would be an important step in understanding the mechanisms of the mitotic checkpoint. Attempts to crystallize Mad2 in complex with several Sgo fragments are presented. These experiments were complemented by NMR and ITC binding studies between the two proteins. A new hypothesis about their interaction and their role in the sister chromatid separation process is supported by these findings. It appears that shugoshin binds Mad2 just like Mad1 and Cdc20 do at the same binding site of the so called *safty belt.*

Since CDK2 and its deregulation play a key role in cancer development much research has focused on finding compounds that can specifically inhibit this kinase. To determine the structural basis of the binding mechanisms of a new small molecule inhibitor for CDK2, crystallization of its complex with CDK2 was attempted. While the crystallization of CDK2 was successful, the structural analysis did not show the compound.

As CDK2 and its action are already well examined and a number of its potential binding partners are known, studying CDK2 – inhibitor interaction was chosen to prove a novel NMR-based drug screening method. This method monitors protein binding via a simple 1D proton NMR signal of the NH side chain ($^{N}H^{\varepsilon}$) of tryptophans and permits rapid screening of potential ligands. The method

also provides quantitative characterization of the antagonist-protein interactions in the form of K_Ds.

7 Zusammenfassung

Diese Dissertation stellt eine Sammlung dreier verschiedener Projekte dar, die sich mit der biochemischen und strukturbiologischen Charakterisierung von Proteinen befassen, die in der Krebsentstehnug eine Rolle spielen. Die große Bandbreite der noch offenen Fragen in der Krebsforschung spiegelt sich auch in den sehr unterschiedlichen Projekten wieder.

Die Kombination von Röntgenkristallographie und NMR Spektroskopie könnte strukturbiologische Hinweise liefern um ein Modell für die Aktivierung von IGF durch seinen Rezeptor IGF-1R zu entwickeln. Um die Kristallstruktur aufzuklären wurden ausführliche Versuche mit verschiedenen Fragmenten unterschiedlicher Länge und mit verschiedenen Mutationen des IGF-1R Rezeptors unternommen. Alle diese Experimente und die dazugehörigen Expressions- und Reinigungsschritte, sowie die NMR Untersuchungen wurden detailiert beschrieben um zukünftige Untersuchungen zu unterstützen.

Die Aufklärung der dreidimensionalen Struktur des Komplexes zwischen Shugoshin und Mad2 wäre ein entscheidender Schritt auf dem Weg zum Verständnis der Mechanismen der Zellteilung. In dieser Arbeit wurden ausführliche Versuche unternommen um Mad2 im Komplex mit verschiedenen Fragmenten des Sgo zu kristallisieren. Diese Experimente wurden durch NMR- und ITC- Bindungsstudien der beiden Proteine ergänzt. Eine neue Hypothese zu ihrer Interaktion und ihrer Rolle bei der Trennung der Schwester-Chromatiden wird durch diese Ergebnisse unterstützt. Offenbar bindet Shugoshin an Mad2 in einer ähnlichen Weise wie auch Mad1 und Cdc20 und zwar in der selben Bindungsstelle des sogenannten *safty belt.*

Da CDK2 und seine Deregulierung eine wichtige Rolle bei der Entstehung von Krebs spielt, hat die Forschung schon viel Zeit darauf verwendet spezifische Inhibitoren dieser Kinase zu finden. Um die strukturelle Grundlage der Bindung eines neuen Inhibitors von CDK2 aufzuklären wurde versucht den Komplex dieses Moleküls mit CDK2 zu kristallisiern. In der gelösten Kristallstruktur konnte der Inhibitor allerdings nicht gefunden werden.

Da CDK2 und seine Wirkungsweise bereits gut verstanden sind und auch etliche mögliche Bindungspartner bekannt sind, wurde die CDK2-Inhibitor Interaktion verwendet um eine neue NMR Screening-Methode zu verifizieren. Bei dieser Methode werden Bindungsstudien an Proteinen mit einem einfachen eindimensionalen NMR Spektrum durchgeführt, wobei das Signal des NHs aus der Tryptophan Seitenkette ($^{N}H^{\varepsilon}$) beobachtet wird. Dabei sind auch quantitative Aussagen über den K_D der Antagonist-Protein Interaktion möglich.

8 Bibliography

Achenbach, T.V., Muller, R. & Slater, E.P. (2000) Bcl-2 independence of flavopiridol-induced apoptosis. J. Biol. Chem. *275,* 32089-97.

Adams, M.J., Blundell, T.L., Dodson, E.J., Dodson, G.G., Vijayan, M., Baker, E.N., Harding, M.M., Hodgkin, D.C., Rimmer, B. & Sheat, S. (1969) Structure of rhombohedral 2 zinc insulin crystals. Nature *224*, 491.

Adams, T.E., Epa, V.C. & Garrett, T.P.J. (2000) Structure of the type I insulin-like growth factor receptor. Cell. Mol. Life Sci. *57*, 1050–93.

Agarwal, M.L., Taylor, W.R., Chernov, M.V., Chernova, O.B. & Stark, G.R. (1998) The p53 network. J. Biol. Chem. 273, 1-4.

Akiyama, T., Yoshida, T., Tsujita, T., Shimizu, M., Mizukami, T., Okabe, M. & Akinaga, S. (1997) G1 phase accumulation induced by UCN-01 is associated with dephosphorylation of Rb and CDK2 proteins as well as induction of CDK inhibitor p21/Cip1/WAF1/Sdi1 in p53-mutated human epidermoid carcinoma A431 cells. Cancer Res. *57*, 1495-1501.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. (2002) Molecular biology of the cell. Garland Science, New York.

Allan, G.J., Tonner, E., Szymanowska, M., Shand, J.H., Kelly, S.M., Phillips, K., Clegg, R.A., Gow, I.F., Beattie, J. & Flint, D.J. (2006) Cumulative mutagenesis of the basic residues in the 201-218 region of insulin-like growth factor-binding protein 5 results in progressive loss of both IGF-1 binding and inhibition of IGF-1 biological actions. Endocrinology *147*, 338-349.

Aramini, J.M., Rossi, P., Anklin, C., Rong Xiao, R. & Montelione, G.T. (2007) Microgram-scale protein structure determination by NMR. Nature Meth. *6*, 491-493.

Bach, L.A., Headey, S.J. & Norton, R.S. (2005) IGF-binding proteins-the pieces are falling into place. Trends Endocrinol. Metab. *16*, 228-34.

Ballard, F.J., Wallace, J.C., Francis, G.L., Read, L.C. & Tomas, F.M. (1996) Des (1-3)-IGF-1: a truncated form of insulin-like growth factor-I. Int. J. Biochem. Cell. Biol. *28*, 1085-7.

Baserga, R. (1995) The insulin-like growth factor I receptor: a key to tumor growth? Cancer Res. *55*, 249-52.

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.

Beck, K.D., Knusel, B. & Hefti, F. (1993) The nature of the trophic action of brain-derived neurotrophic factor des(1-3)-insulin-like growth factor-1, and basic fibroblast

growth factor on mesencephalic dopaminergic neurons developing in culture. Neurosci. *52*, 855-66.

Benson, C., White, J., De Bono, J., O'Donnell, A., Raynaud, F., Cruickshank, C., McGrath, H., Walton, M., Workman, P., Kaye, S., Cassidy, J., Gianella-Borradori, A., Judson, I. & Twelves, C. (2007) A phase I trial of the selective oral cylin-dependent kinase inhibitor seliciclib (CYC202, R-roscovitine) administered twice daily for 7 days every 21 days. Br. J. Cancer. *96*, 29-37.

Bernards, R. (2008) Cancer: Entangled pathways. Nature.*455*, 479-80.

Boos, D., Kuffer, C., Lenobel, R., Körner, R. & Stemmann, O. (2008) Phosphorylation-dependend binding of cyclin B1 to a Cdc6-like domain of human separase. J. Biol. Chem. *283*, 816-823.

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.

Buehler, C., Dreessen, J., Mueller, K., So, P.T., Schilb, A., Hassiepen, U., Stoeckli, K.A. & Auer, M. (2005) Multi-photon excitation of intrinsic protein fluorescence and its application to pharmaceutical drug screening. Assay Drug Dev. Technol. *3*, 155-167.

Bunch, R.T. & Eastman, A. (1996) Enhancement of cisplatin-induced cytotoxicity by 7-hydroxystaurosporine (UCN-01), a new G2-checkpoint inhibitor. Clin. Cancer Res. *2*, 791-797.

Bunn, R.C. & Fowlkes, J.L. (2003) Insulin-like growth factor binding protein proteolysis. Trends Endoc. & Metabol. *14*, 176-181.

Cahill, D.P., Lengauer, C., Yu, J., Riggins, G.J., Willson, J.K., Markowitz, S.D., Kinzler, K.W. & Vogelstein, B. (1998) Mutations of mitotic checkpoint genes in human cancers. Nature *392*, 300-3.

Carlsson–Skwirut, C., Jornvall, H., Holmgren, A., Andersson, C., Bergman, T., Lundquist, B. & Sara, V.R. (1986) Isolation and characterisation of variant IGF-I as well as IGF-2 from adult human brain. FEBS Lett. *201*, 46-50.

Carlson-Skwirut, C., Lake, M., Hartmanis, M., Hall, K. & Sara, V.R. (1989) A comparison of the biological activity of the recombinant intact, and truncated insulin-like growth factor 1 (IGF-1). Biochim. Biophys. Acta *1011*, 192-7.

Carrick, F.E., Wallace, J.C. & Forbes, B.E. (2002) The interaction of Insulin-like growth factors (IGFs) with Insulin-like growth factor binding proteins (IGFBPs): a review. *Lett. Pep. Sci. 8*, 147–153.

Carroll, P.V., Christ, E.R., Umpleby, A.M., Gowrie, I., Jackson, N., Bowes, S.B., Hovorka, R., Croos, P., Sönksen, P.H. & Russell-Jones, D.L. (2000) IGF-I treatment in adults with type 1 diabetes: effects on glucose and protein metabolism in the fasting state and during a hyperinsulinemic-euglycemic amino acid clamp. Diabetes *49*, 789-96.

CCP4, Collaborative Computational Project, Number 4. (1994) The CCP4 Suite: Programs for protein crystallography. Acta Cryst. D *50*, 760-763.

Chene, P. (2003) Inhibiting the p53–MDM2 interaction: an important target for cancer therapy. Nat. Rev. Cancer *3*, 102-109.

Chung, C.T. Niemela, S.L. & Miller, R.H. (1989) One-step preparation of competent Escherichia coli: Transformation and storage of bacterial cells in the same solution. Proc. Natl Acad. Sci. USA *86*, 2172-2175.

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

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.

Danley, D.E. (2006) Crystallization to obtain protein-ligand complexes for structure-aided drug design. Acta Crystallogr. D *62*, 569-75.

Daughaday, W. & Salmon Jr, W.D. (1999) Historical perspective: the origins and development of the somatomedin hypothesis, in: The IGF system: Molecular biology, physiology and clinical applications. Rosenfeld, R.G. and Roberts, C.T. (eds), Humana Press, Totowa, New Jersey. 1-15.

Davis, S.T., Benson, B.G. & Bramson, H.N. (2001) Prevention of chemotherapy-induced alopecia in rats by CDK inhibitors. Science *291*,134-137.

De Antoni, A., Pearson, C.G., Cimini, D., Canman, J.C., Sala, V., Nezi, L., Mapelli, M., Sironi, L., Faretta, M., Salmon, E.D. & Musacchio, A. (2005) The Mad1/Mad2 complex as a template for Mad2 activation in the spindle assembly checkpoint. Curr. Biol. *15*, 214-25.

De Bondt, H.L., Rosenblatt, J., Jancarik, J., Jones, H.D., Morgan, D.O. & Kim S.H. (1993) Crystal structure of cyclin-dependent kinase 2. Nature *363*, 595-602.

De Meyts P. (2008) The insulin receptor: a prototype for dimeric, allosteric membrane receptors? Trends Biochem Sci., *33*, 376-84.

De Meyts P. & Whittaker J. (2002) Structural biology of insulin and IGF1 receptors: implications for drug design. Nat Rev Drug Discov. *10*, 769-83.

Dodatko. T., Fedorov, A.A., Grynberg, M., Patskovsky, Y., Rozwarski, D.A., Jaroszewski, L., Aronoff-Spencer, E., Kondraskina, E., Irving, T., Godzik, A. & Almo, S.C. (2004) Crystal structure of the actin binding domain of the cyclase-associated protein. Biochemistry *43*,10628-41.

Draviam, V.M., Xie, S. & Sorger, P.K. (2004) Chromosome segregation and genomic stability. Curr. Opin. Genet. Dev. *14*, 120-5.

Duan, C. (1997) The insulin-like growth factor system and its biological actions in fish. Am. Zool. *37*, 489-501.

Duan C., Ding J., Li Q., Tsai W. & Pozios K. (1999) Insulin-like growth factor binding protein 2 is a growth inhibitory protein conserved in zebrafish. Proc Natl Acad Sci USA *96*, 15274-9.

Easton, J., Wei, T., Lahti, J.M. & Kidd, V.J. (1998) Disruption of the cyclin D/cyclin-dependent kinase/INK4/retinoblastoma protein regulatory pathway in human neuroblastoma. Cancer Res. *58*, 2624-32.

Evans, T., Rosenthal, E.T., Youngblom, J., Distel, D. & Hunt, T. (1983) Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. Cell *33*, 389-96.

Ewen, M.E. (1994) The cell cycle and the retinoblastoma protein family. Cancer Metastasis Rev. *13*, 45–66.

Fernandez-Tornero, C., Lozano, R.M., Rivas, G., Jimenez, M.A., Ständker, L., Diaz-Gonzalez, D., Forssmann, W.G., Cuevas, P., Romero, A. & Giménez-Gallego, G. (2005) Synthesis of the blood circulating C-terminal fragment of insulin-like growth factor binding protein 4 in its native conformation. Crystallization, heparin and IGF binding, and osteogenic activity. J. Biol. Chem. *280*, 18899–18907.

Firestein, R., Bass, A.J., Kim, S.Y., Dunn, I.F., Silver, S.J., Guney, I., Freed, E., Ligon, A.H., Vena, N., Ogino, S., Chheda, M,G., Tamayo, P., Finn, S., Shrestha, Y., Boehm, J.S., Jain, S., Bojarski, E., Mermel, C., Barretina, J., Chan, J.A., Baselga, J., Tabernero, J., Root, D.E., Fuchs, C.S., Loda, M., Shivdasani, R.A., Meyerson, M. & Hahn W.C. (2008) CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity. Nature *455*, 547-51.

Firth, S.M. & Baxter, R.C. (2002) Cellular actions of the insulin-like growth factor binding proteins. Endocr. Rev. *23*, 824-854.

Firth, S.M., Clemmons, D.R. & Baxter, R.C. (2001) Mutagenesis of basic amino acids in the carboxy-terminal region of insulin-like growth factor binding protein-5 affects acid-labile subunit binding. Endocrinology *14*, 2147-2150.

Francis, G.L., Upton, F.M., Ballard, F.J., McNeil, K.A. & Wallace, J.C. (1988) Insulin-like growth factors 1 and 2 in bovine colostrum. Sequences and biological activities compared with those of a potent truncated form. Biochemical J. *251*, 95-103.

Furstenberger, G. & Senn, H.J. (2002) Insulin-like growth factors and cancer. Lancet Oncol. *3*, 298.

Garrett, T.P., McKern, N.M., Lou, M., Frenkel, M.J., Bentley, J.D., Lovrecz, G.O., Elleman, T.C., Cosgrove, L.J. & Ward, C.W. (1998) Crystal structure of the first three domains of the type-1 insulin-like growth factor receptor. Nature *394*, 395-9.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D. & Bairoch, A. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. *31*, 3784-8.

Gill, S.C. & von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. *182*, 319-26.

Golovanov, A.P., Blankley, R.T., Avis, J.M. & Bermel, W. (2007) Isotopically discriminated NMR spectroscopy: A tool for investigating complex protein interactions in vitro. J. Am. Chem. Soc. *129*, 6528-6535.

Gray, D., Subramanian, S. (2001) Choice of cellular protein expression system. Curr. Protoc. Protein Sci. 5.16.1-5.16.34.

Gruber, S., Haering, C.H. & Nasmyth, K. (2003) Chromosomal cohesin forms a ring. Cell *112*, 765-77.

Guan, J., Bennet, L., Gluckman, P.D. & Gunn, A.J. (2003) Insulin-like growth factor-1 and post-ischemic brain injury. Prog. Neurobiol. *70*, 443-62.

Gyuris, J., Golemis, E., Chertkov, H. & Brent, R. (1993) Cdi1, a human G1 and S phase protein phosphatase that associates with CDK2. Cell *75*, 791–803.

Haering, C.H. & Nasmyth, K. (2003) Building and breaking bridges between sister chromatids. Bioessays *25*, 1178-91.

Halloran, B.P. & Spencer, E.M. (1988) Dietary phosphorus and 1,2-dihydroxyvitamin D metabolism: influence of insulin-like growth factor I. Endocrinology *123*, 1225-9.

Hamid, T. & Kakar, S.S. (2003) PTTG and cancer. Histol. Histopathol. *18*, 245-51.

Hannon, G.J. & Beach, D. (1994) p15INK4B is a potential effector of TGFβ-induced cell cycle arrest. Nature *371*, 257-261.

Hardcastle, I.R., Golding, B.T., Griffin, R.J. (2002) Designing inhibitors of cyclin-dependent kinases. Annu. Rev. Pharmacol. Toxicol. *42*, 325-48.

Harper, J.W., Adami, G.R., Wei, N., Keyomarsi, K. & Elledge, S.J. (1993) The p21 CDK-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell *75,* 805-816.

Hatakeyama, M., Brill, J.A., Fink, G.R. & Weinberg, R.A. (1994) Collaboration of G1 cyclins in the functional inactivation of the retinoblastoma protein. Genes Dev. *8*, 1759-1771.

Hauf, S., Roitinger, E., Koch, B., Dittrich, C.M., Mechtler, K. & Peters, J.M. (2005) Dissociation of cohesin from chromosome arms and loss of arm cohesion during early mitosis depends on phosphorylation of SA2. PloS. Biol. *3*, e69.

Headey, S.J., Leeding, K.S., Norton, R.S. & Bach, L.A. (2004) Contributions of the N- and C-terminal domains of IGF binding protein-6 to IGF binding. J. Mol. Endocr. *33*, 377-86.

Hellawell, G.O., Turner, G.D., Davies, D.R., Poulsom, R., Brewster, S.F. & Macaulay, V.M. (2002) Expression of the type 1 insulin-like growth factor receptor is up-regulated in primary prostate cancer and commonly persists in metastatic disease. Cancer Res. *62*, 2942-50.

Hermeking, H., Lengauer, C., Polyak, K., He, T.C., Zhang, L., Thiagalingam, S., Kinzler, K.W. & Vogelstein B. (1997) 14-3-3 sigma is a p53-regulated inhibitor of G2/M progression. Mol. Cell. *1*, 3-11.

Holland, A.J, Böttger, F., Stemmann, O. & Taylor, S.S. (2007) Protein phosphatase 2A and separase form a complex regulated by separase autocleavage. J. Biol. Chem. *282*, 24623-32.

Humbel, R.E. (1990) Insulin-like growth factor-I, and II. Eur. J. Biochem. *190*, 445-62.

Huwe, A., Mazitschek, R. & Giannis, A. (2003) Small molecules as inhibitors of cyclin-dependent kinases. Angew. Chem. Int. Ed. Engl. *42*, 2122-38.

Ivanov, D. & Nasmyth, K. (2005) A topological interaction between cohesin rings and a circular minichromosome. Cell *122*, 849-60.

Jahnke, W. & Widmer, H. (2004) Protein NMR in biomedical research. Cell. Mol. Life Sci. *61*, 580-599.

Jallepalli, P.V. & Lengauer, C. (2001) Chromosome segregation and cancer: cutting through the mystery. Nat. Rev. Cancer *1*, 109-17.

Jallepalli, P.V., Waizenegger, I.C., Bunz, F., Langer, S., Speicher, M.R., Peters, J.M., Kinzler, K.W., Vogelstein, B. & Lengauer, C. (2001) Securin is required for chromosomal stability in human cells. Cell *105*, 445-57.

Jayaraman, L. & Prives, C. (1999) Covalent and noncovalent modifiers of the p53 protein. Cell Mol. Life Sci. *55*, 76-87.

Jeffrey, P.D., Russo, A.A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J. & Pavletich, N.P. (1995) Mechanism of CDK activation revealed by the structure of a cyclin A-CDK2 complex. Nature *376*, 313-320.

Jentsch, S. (2007) Personal communication.

Jones, J.L. & Clemmons, D.R. (1995) Insulin-like growth factors and their binding proteins: biological actions. Endocr. Rev. *16*, 3-34.

Jung, G.W., Spencer, E.M. & Lue, T.F. (1998) Growth hormone enhances regeneration of nitric oxide synthase-containing penile nerves after cavernous nerve neurotomy in rats. Journal of Urology *160*, 1899-904.

Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst. *26*, 795-800.

Kaspar, B.K., Lladó, J., Sherkat, N., Rothstein, J.D. & Gage, F.H. (2003) Retrograde viral delivery of IGF-1 prolongs survival in a mouse ALS model. Science *301*,839-42.

Kato, J., Matsushime, H., Hiebert, S.W., Ewen, M.E. & Sherr, C.J. (1993) Direct binding of cyclin D to the retinoblastoma gene product (pRb) and pRb phosphorylation by the cyclin D-dependent kinase CDK4. Genes Dev. *7*, 331-342.

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, 215-44.

Kibbey, M.M., Jameson, M.J., Eaton, E.M. & Rosenzweig, S.A. (2006) Insulin-like growth factor binding protein-2: Contributions of the C-terminal domain to insulin-like growth factor-1 binding. Mol. Pharm. *69*, 833-845.

Kim, B., Cheng, H.L., Margolis, B. & Feldman, E.L. (1998) Insulin receptor substrate 2 and Shc play different roles in insulin-like growth factor-I signaling. J. Biol. Chem. *273*, 34543-50.

Kim, J.H., Kang, M.J., Park, C.U., Kwak, H.J., Hwang, Y. & Koh, G.Y. (1999) Amplified CDK2 and cdc2 activities in primary colorectal carcinoma. Cancer *85*, 546-553.

Kitajima, T.S., Hauf, S., Ohsugi, M., Yamamoto, T. & Watanabe, Y. (2005) Human Bub1 defines the persistent cohesion site along the mitotic chromosome by affecting Shugoshin localization. Curr. Biol. *15*(4), 353-9.

Kitajima, T.S., Kawashima, S.A. & Watanabe, Y. (2004) The conserved kinetochore protein shugoshin protects centromeric cohesion during meiosis. Nature *427*, 510-7.

Kitajima, T.S., Sakuno, T., Ishiguro, K., Iemura, S., Natsume, T., Kawashima, S.A. & Watanabe Y. (2006) Shugoshin collaborates with protein phosphatase 2A to protect cohesin. Nature *441*, 46-52.

Klages, J., Coles, M. & Kessler, H. (2006) NMR-based screening: a powerful tool in fragment-based drug discovery. Mol. Biosyst. *2*, 319-331.

Ko, L.J. & Prives, C. (1996) p53: puzzle and paradigm. Genes Dev. *10*,1054-72.

Kobayashi, H., Stewart, E., Poon, R., Adamczewski, J.P., Gannon, J. & Hunt, T. (1992) Identification of the domains in cyclin A required for binding to, and activation of, p34cdc2 and p32cdk2 protein kinase subunits. Mol. Biol. Cell. *3*, 1279-94.

Kops, G.J., Weaver, B.A. & Cleveland, D.W. (2005) On the road to cancer: aneuploidy and the mitotic checkpoint. Nat. Rev. Cancer *5*, 773-85.

Kumar, S.K., Hager, E., Pettit, C., Gurulingappa, H., Davidson, N.E. & Khan, S.R. (2003) Design, synthesis, and evaluation of novel Boronic-chalcone derivatives as antitumor agents. J. Med. Chem. *46*, 2813-2815.

Lapointe, J., Lachance, Y., Labrie, Y. & Labrie, C. (1996) A p18 mutant defective in CDK6 binding in human breast cancer cells. Cancer Res. *56*, 4586-9.

Lees, E.M. & Harlow, E. (1993) Sequences within the conserved cyclin box of human cyclin A are sufficient for binding to and activation of cdc2 kinase. Mol. Cell. Biol. *13*, 1194-201.

Lepre, C.A., Moore, J.M. & Peng, J.W. (2004) Theory and applications of NMR-based screening in pharmaceutical research. Chem. Rev. *104*, 3641-3676.

LeRoith, D., Werner, H., Faria, T.N., Kato, H., Adamo, M. & Roberts, C.T. (1993) Insulin-like growth factor receptor: implications for nervous function. Ann. NY Acad. Sci. *692*, 22-32.

LeRoith, D., Werner, H., Beitner-Johnson, D. & Roberts, C.T. (1995) Molecular and cellular aspects of the insulin-like growth factor I receptor. Endocr. Rev. *16*, 143-63.

LeRoith, D. & Roberts, C.T. (2003) The insulin-like growth factor system and cancer. Cancer Lett. *195*, 127-37.

LeRoith, D. & Helman, L. (2004) The new kid on the block(ade) of the IGF-1 receptor. Cancer Cell *5*, 201-2.

Leslie, A.G.W. (1991) Molecular data processing. Crystallographic Computing. Oxford University Press, 50–61.

Leventhal, P.S., Russel, J.W. & Feldman, E.L. (1999) IGFs and the nervous system. In: The IGF System. R.G. Rosenfeld and C. Roberts (eds). Humana Press, Totowa, New Jersey.

Levine, A.J. (1997) p53, the cellular gatekeeper for growth and division. Cell *88*, 323-331.

Longo, V.D. & Finch, C.E. (2003) Evolutionary medicine: from dwarf model systems to healthy centenarians? Science *299*, 1342-6.

Lou, M., Garrett, T.P., McKern, N.M., Hoyne, P.A., Epa, V.C., Bentley, J.D., Lovrecz, G.O., Cosgrove, L.J., Frenkel, M.J. & Ward, C.W.I. (2006) Crystal structure of the first three domains of the human insulin receptor reveals major differences from the IGF-1 receptor in the regions governing ligand specificity. Proc. Natl Acad. Sci. USA *103*, 12429-34.

Majumdar, S. (2005) personal communication

Makrides, S.C. (1996) Strategies for achieving high-level expression of genes in Escherichia coli. Microbiol. Rev. *60*, 512-538.

Mapelli, M., Massimiliano, L., Santaguida, S. & Musacchio, A. (2007) The Mad2 conformational dimer: structure and implications for the spindle assembly checkpoint. Cell *131*, 730-43.

Mapelli, M. & Musacchio, A. (2007) MAD contortions: conformational dimerization boosts spindle checkpoint signaling. Curr. Opin. Struct. Biol. *17*, 716-25.

Margolis, R.L. (2005) Tetraploidy and tumor development. Cancer Cell *8*, 353-4.
Marx, J. (2002) Debate surges over the origins of genomic defects in cancer. Science *297*, 544-6.

Matthews, B.W. (1968) Solvent content of protein crystals. J. Mol. Biol. *33*, 491–497.

Matthews, B.W. (1974) Determination of molecular weight from protein crystals. J. Mol. Biol. *82*, 513-26.

Maures, T.J. & Duan, C. (2002) Structure, developmental expression, and physiological regulation of zebrafish IGF-binding protein-1. Endocrinology *143*: 1858-1871.

McClue, S.J., Blake, D., Clarke, R., Cowan, A., Cummings, L., Fischer, P.M., MacKenzie, M., Melville, J., Stewart, K., Wang, S., Zhelev, N., Zheleva, D. & Lane, D.P. (2002) In vitro and in vivo antitumor properties of the cyclin dependent kinase inhibitor CYC202 (R-roscovitine). Int. J. Cancer *102*, 463-468.

McDonald, E.R. & el-Deiry, W.S. (2000) Cell cycle control as a basis for cancer drug development. Int. J. Oncol. *16*, 871-86.

McGuinness, B.E., Hirota, T., Kudo, N.R., Peters, J.M. & Nasmyth, K. (2005) Shugoshin prevents dissociation of cohesin from centromeres during mitosis in vertebrate cells. PloS. Biol. *3*, e86.

McKern, N.M., Lawrence, M.C., Streltsov, V.A., Lou, M.Z., Adams, T.E., Lovrecz, G.O., Elleman, T.C., Richards, K.M., Bentley, J.D., Pilling, P.A., Hoyne, P.A., Cartledge, K.A., Pham, T.M., Lewis, J.L., Sankovich, S.E., Stoichevska, V., Da Silva, E., Robinson, C.P., Frenkel, M.J., Sparrow, L.G., Fernley, R.T., Epa, V.C. & Ward, C.W. (2006) Structure of the insulin receptor ectodomain reveals a folded-over conformation. Nature *443*, 218-21.

McKern, N.M., Lou, M., Frenkel, M.J., Verkuylen, A., Bentley, J.D., Lovrecz, G.O., Ivancic, N., Elleman, T.C., Garrett, T.P., Cosgrove, L.J. & Ward, C.W. (1997) Crystallization of the first three domains of the human insulin-like growth factor-1 receptor. Protein Sci. *6*, 2663-2666.

Meijer, L. & Raymond, E. (2003) Roscovitine and other purines as kinase inhibitors. From starfish oocytes to clinical trials. Acc. Chem. Res. *36*, 417-425.

Meijer, L., Borgne, A., Mulner, O., Chong, J.P., Blow, J.J., Inagaki, N., Inagaki, M., Delcros, J.G. & Moulinoux, J.P. (1997) Biochemical and cellular effects of roscovitine, a potent and selective inhibitor of the cyclin-dependent kinases cdc2, cdk2 and cdk5. Eur. J. Biochem. *243*, 527-536.

Meijer, L., Jezequel, A. & Ducommun, B. (2000) Progress in cell cycle research, vol. 4. Plenum Press, New York.

Meyer, B. & Peters, T. (2003) NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. Angew. Chem. Int. Ed. *42*, 864-90.

Michel, L.S., Liberal, V., Chatterjee, A., Kirchwegger, R., Pasche, B., Gerald, W., Dobles, M., Sorger, P.K., Murty, V.V. & Benezra, R. (2001) MAD2 haplo-insufficiency causes premature anaphase and chromosome instability in mammalian cells. Nature *409*, 355-9.

Miyakoshi, N., Richman, C., Kasukawa, Y., Linkhart, T.A., Baylink, D.J. & Mohan, S. (2001) Evidence that IGF-binding protein-5 functions as a growth factor. J. Clin. Invest. *107*, 73-81.

Miyazaki, W.Y. & Orr-Weaver, T.L. (1994) Sister-chromatid cohesion in mitosis and meiosis. Annu. Rev. Genet. *28*, 167-87.

Morgan, D.O. & De-Bondt, H.L. (1994) Protein kinase regulation: insights from crystal structure analysis. Curr. Opin. Cell. Biol. *6*, 239-246.

Motwani, M., Li, X. & Schwartz, G.K. (2000) Flavopiridol, a cyclin-dependent kinase inhibitor, prevents spindle inhibitor-induced endoreduplication in human cancer cells. Clin. Cancer Res. *6*, 924-932.

Musacchio, A. & Salmon, E.D. (2007) The spindle-assembly checkpoint in space and time. Nature Rev. Mol. Cell Biol. *8*, 379-393.

Nakae, J., Kido, Y. & Accili, D. (2001) Distinct and overlapping functions of insulin and IGF-I receptors. Endocr. Rev. *22*, 818-35.

Nasmyth, K. (2005) How do so few control so many? Cell *120*, 739-46.

Nasmyth, K. & Haering, C.H. (2005) The structure and function of SMC and kleisin complexes. Annu. Rev. Biochem. *74*, 595-648.

Nef, S., Verma-Kurvari, S., Merenmies, J., Vassalli, J.D., Efstratiadis, A., Accili, D. & Parada, L.F. (2003) Testis determination requires insulin receptor family function in mice. Nature *426*: 291-295.

Norby, F.L., Wold, L.E., Duan, J., Hintz, K.K. & Ren, J. (2002) IGF-I attenuates diabetes-induced cardiac contractile dysfunction in ventricular myocytes. Am. J. Physiol. Endocrinol. Metab. *283*, 658-66.

Payet, L.D., Wang, X.H., Baxter, R.C. & Firth, M. (2003) Amino- and carboxy-terminal fragments of insulin-like growth factor (IGF) binding protein-3 cooperate to bind IGFs with high affinity and inhibit IGF receptor interactions. Endocrinology *144*, 2797-806.

Peng, J.W., Moore, J. & Abdul-Manan, N. (2004) NMR experiments for lead generation in drug discovery. Prog. NMR Spectr. *44*, 225-256.

Pereira, J.J., Meyer, T., Docherty, S.E., Reid, H.H., Marshall, J., Thompson, E.W., Rossjohn, J. & Price, J.T. (2004) Bimolecular interaction of insulin-like growth factor (IGF) binding protein-2 with alpha-beta3 negatively modulates IGF-I-mediated migration and tumor growth. Cancer Res. *64*, 977-84.

Perks, C.M., Newcomb, P.V., Norman, M.R. & Holly, J.M. (1999) Effect of insulin-like growth factor binding protein-1 on integrin signaling and the induction of apoptosis in human breast cancer cells. J. Mol. Endocrinol. *22*, 141–50.

Pihan, G. & Doxsey, S.J. (2003) Mutations and aneuploidy: co-conspirators in cancer? Cancer Cell *4*, 89-94.

Pines. J. & Hunt, T. (1987) Molecular cloning and characterization of the mRNA for cyclin from sea urchin eggs. EMBO J. *10*, 2987-95.

Pollak, M.N., Schernhammer, E.S. & Hankinson, S.E. (2004) Insulin-like growth factors and neoplasia. Nature Rev. Cancer *4*, 505-18.

Rabitsch, K.P., Gregan, J., Schleiffer, A., Javerzat, J.P., Eisenhaber, F. & Nasmyth, K. (2004) Two fission yeast homologs of Drosophila Mei-S332 are required for chromosome segregation during meiosis I and II. Curr. Biol. *14*, 287-301.

Ramaswamy, S., Ross, K.N., Lander, E.S. & Golub, T.R. (2003) A molecular signature of metastasis in primary solid tumors. Nat. Genet. *33*, 49-54.

Raynaud, F.I., Whittaker, S.R., Fischer, P.M., McClue, S., Walton, M.I., Barrie, S.E., Garrett, M.D., Rogers, P., Clarke, S.J., Kelland, L.R., Valenti, M., Brunton, L., Eccles, S., Lane, D.P. & Workman, P. (2005) In vitro and in vivo pharmacokinetic-pharmacodynamic relationships for the trisubstituted aminopurine cyclin-dependent kinase inhibitors olomoucine, bohemine and CYC202. Clin. Cancer Res. *11*, 4875-87.

Rechler, M.M. (1985) The nature and regulation of the receptors for insulin-like growth factors. Ann. Rev. Physiol. *47*, 425-42.

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

Ricort, J.M. (2004) Insulin-like growth factor binding protein (IGFBP) signalling. Growth hormone & IGF research *14*, 277-86.

Rieder, C.L., Schultz, A., Cole, R. & Sluder, G. (1994) Anaphase onset in vertebrate somatic cells is controlled by a checkpoint that monitors sister kinetochore attachment to the spindle. J. Cell. Biol. *127*, 1301-10.

Robey, R.W., Medina-Perez W.Y. & Nishiyama, K. (2001) Overexpression of the ATP-binding cassette half-transporter, ABCG2 (MXR/BCRP/ABCP1), in flavopiridol-resistant human breast cancer cells. Clin. Cancer Res. *7*, 145-152.

Rothweiler, U. (2006) personal communication.

Rothweiler, U., Czarna, A., Weber, L., Popowicz, G.M., Brongel, K., Kowalska, K., Orth, M., Stemmann, O. & Holak TA. (2008) NMR screening for lead compounds using tryptophan-mutated proteins. J. Med. Chem. *51*, 5035-42.

Russo, A., Jeffrey, P.D. & Pavletich, N.P. (1996) Structural basis of cyclin dependent kinase activation by phosphorylation. Nat. Struct. Biol. *3*, 696-700.

Sanchez, Y., Wong, C., Thoma, R.S., Richman, R., Wu, Z., Piwnica-Worms, H. & Elledge, S.J. (1997) Conservation of the Chk1 checkpoint pathway in mammals: linkage of DNA damage to Cdk regulation through Cdc25. Science *277*, 1497-501.

Santamaria, D., Barriere, C., Cerqueira, A., Hunt, S., Tardy, C., Newton, K., Caceres, J.F., Dubus, P., Malumbres, M. & Barbacid M. (2007) Cdk1 is sufficient to drive the mammalian cell cycle. Nature *448*, 811-5.

Sara, V.R., Carlsson-Skwirut, C., Anderson, C., Hall, E., Sjorgen, B., Holmgren, A. & Jornvall, H. (1986) Characterization of somatomedins from human fetal brain: identification of a variant form of insulin-like growth factor I. Proc. Natl. Acad. Sci. USA *83*, 4904-7.

Sato, H., Nishimura, S., Ohkubo, T., Kyoguku, Y., Koyama, S. & Kobayashi, Y. (1993) Three-dimensional structure of human insulin-like growth factor I (IGF-I) determined by 1H-NMR and distance geometry, Int. J. Pept. Protein Res. *41*, 433 40.

Sausville, E.A., Arbuck, S.G., Messmann, R., Headlee, D., Bauer, K.S., Lush, R.M., Murgo, A., Figg, W.D., Lahusen, T., Jaken, S., Jing, X., Roberge, M., Fuse, E., Kuwabara, T. & Senderowicz, A.M. (2001) Phase I trial of 72-hour continuous infusion UCN-01 in patients with refractory neoplasms. J. Clin. Oncol. *19*, 2319-33.

Scanlan, M.J., Gout, I., Gordon, C.M., Williamson, B., Stockert, E., Gure, A.O., Jäger, D., Chen, Y.T., Mackay, A., O'Hare, M.J. & Old, L.J. (2001) Humoral immunity to human breast cancer: antigen definition and quantitative analysis of mRNA expression. Cancer Immun. *1*, 4.

Schäffer, L., Brissette, R.E., Spetzler, J.C., Pillutla, R.C., Østergaard, S., Lennick, M., Brandt, J., Fletcher, P.W., Danielsen, G.M,, Hsiao, K.C., Andersen, A.S., Dedova, O., Ribel, U., Hoeg-Jensen, T., Hansen, P.H., Blume, A.J., Markussen, J. & Goldstein, N,I. (2003) Assembly of high-affinity insulin receptor agonists and antagonists from peptide building blocks. Proc. Natl Acad. Sci. USA *100*, 4435–-39.

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*, 368-79.

Schlechter, N.L., Russell, S.M., Greenberg, S., Spencer, E.M. & Nicoll, C.S. (1986) A direct growth effect of growth hormone in rat hindlimb shown by arterial infusion. Am. J. Physiol. *250*, 231-5.

Senderowicz, A.M. (1999) Flavopiridol: the first cyclin-dependent kinase inhibitor in human clinical trials. Invest. New Drugs *17*, 313-320.

Senderowicz, A.M. (2003) Novel small molecule cyclin-dependent kinases modulators in human clinical trials. Cancer Biol. Ther. *2*, Suppl.1, S84-S96.

Shah, J.V. & Cleveland, D.W. (2000) Waiting for anaphase: Mad2 and the spindle assembly checkpoint. Cell *103*, 997-1000.

Shapiro, G.I. (2006) Cyclin-dependent kinase pathways as targets for cancer treatment. J. Clin. Oncol. *24*, 1770-1783.

Sherr, C. J. (1996) Cancer cell cycles. Science *274*, 1672-77.

Shi, Y., Zou, M., Farid, N.R. & al-Sedairy, S.T. (1996) Evidence of gene deletion of p21 (WAF1/CIP1), a cyclin-dependent protein kinase inhibitor, in thyroid carcinomas. Br. J. Cancer *74*, 1336-41.

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.

Sironi, L., Mapelli, M., Knapp, S., De Antoni, A., Jeang, K.T. & Musacchio, A. (2002) Crystal structure of the tetrameric Mad1-Mad2 core complex: implications of a 'safety belt' binding mechanism for the spindle checkpoint. EMBO J. *21*, 2496-506.

Sitar, T., Popowicz, G.M., Siwanowicz, I., Huber, R. & Holak, T.A. (2006) Structural basis for the inhibition of insulin-like growth factors by insulin-like growth factor-binding proteins. Proc. Natl Acad. Sci. USA *103*, 13028-33.

Sitar, T. (2007) Personal communication.

Siwanowicz, I., Popowicz, G.M., Wisniewska, M., Huber, H., Kuenkele, K.P.Lang, K. & Holak, T.A. (2005) Structural basis for the regulation of insulin-like growth factors by IGF binding proteins. Structure *13*, 155-167.

Stemmann, O. (2006) personal communication

Stemmann, O. (2008) personal communication

Stockman, B.J. & Dalvit, C. (2002) NMR screening techniques in drug discovery and drug design. Prog. NMR Spectr. *41*, 187-231.

Stoll, R., Renner, C., Hansen, S., Palme, S., Klein, C., Belling, A., Zeslawski, W., Kamionka, M., Rehm, T., Muhlhahn, 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.

Suh, D.Y., Hunt, T.K. & Spencer E.M. (1992) Insulin-like growth factor-I reverses the impairment of wound healing induced by corticosteroids in rats. Endocrinology *131*, 2399-406.

Sumara, I., Vorlaufer, E., Stukenberg, P.T., Kelm, O., Redemann, N., Nigg, E.A. & Peters, J.M. (2002) The dissociation of cohesin from chromosomes in prophase is regulated by Polo-like kinase. Mol. Cell. *9*, 515-25.

Tang, Z., Sun, Y., Harley, S. E., Zou, H. & Yu, H. (2004) Human Bub1 protects centromeric sister-chromatid cohesion through Shugoshin during mitosis. Proc. Natl. Acad. Sci. USA *101*, 18012-7.

Taylor, W.R. & Stark, G.R. (2001) Regulation of the G2/M transition by p53. Oncogene *20*, 1803-15.

Thissen, J. P., Ketelslegers, J. M. & Underwood, L. E. (1994) Nutritional regulation of the insulin-like growth factors. Endo. Rev*. 15,* 80-101 (1994).

Tsumoto, K., Ejima, D., Kumagai, I. & Arakawa, T. (2003) Practical considerations in refolding proteins from inclusion bodies. Protein Expr. Purif. *28*, 1-8.

Uhlmann, F., Wernic, D., Poupart, M.A., Koonin, E.V. & Nasmyth, K. (2000) Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast. Cell *103*, 375-86

Vajdos, F.F., Ultsch, M., Schaffer, M.L., Deshayes, K.D., Liu, J., Skelton, N.J. & de Vos, A.M. (2001) Crystal structure of human insulin-like growth factor-1: detergent binding inhibits binding protein interactions. Biochemistry *40,* 11022-29.

Vassilev, L.T. (2006) MDM2 inhibitors for cancer therapy. Trends Mol. Med. *13*, 23-31.

Vilenchik, L.Z., Griffith, J.P. & St. Clair, N. (1998) Protein crystals as novel microporous materials. J. Am. Chem. Soc. *120*, 4290-94 .

Vousden, K.H. & Lane, D.P. (2007) p53 in health and disease. Nature Rev. Mol. Cell Biol. *8*, 275-283.

Wadler, S. (2001) Perspectives for cancer therapies with CDK2 inhibitors. Drug. Resist. Updat. *4*, 347-67.

Wang, Q., Fan, S., Eastman, A., Worland, P.J., Sausville, E.A. & O'Connor, P.M. (1996) UCN-01: a potent abrogator of G2 checkpoint function in cancer cells with disrupted p53. J. Natl. Cancer Inst. *88*, 956-65.

Wang, T.L., Rago, C., Silliman, N., Ptak, J., Markowitz, S., Willson, J.K., Parmigiani, G., Kinzler, K.W., Vogelstein, B. & Velculescu, V.E. (2002) Prevalence of somatic alterations in the colorectal cancer cell genome. Proc. Natl. Acad. Sci. USA *99*, 3076-80.

Watanabe, Y.& Kitajima, T.S. (2005) Shugoshin protects cohesin complexes at centromeres. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *360*, 515-21.

Wasch, R. & Engelbert, D. (2005) Anaphase-promoting complex-dependent proteolysis of cell cycle regulators and genomic instability of cancer cells. Oncogene *24*, 1-10.

Weiss, M.A., Frank, B.H., Khait, I., Pekar, A., Heiney, R., Shoelson, S.E. & Neuringer, L.J. (1990) NMR and photo-CIDNP studies of human proinsulin and prohormone processing intermediates with application to endopeptidase recognition. Biochemistry *29*, 8389-401.

Weizenegger, I.C., Hauf, S., Meinke, A. & Peters, J.M. (2000) Two distinct pathways remove mammalian cohesin from chromosme arms in prophase and from centromeres in anaphase. Cell *103*, 399-410.

White, M.F. & Kahn, C.R. (1989) Cascade of autophosphorylation in the beta-subunit of the insulin receptor. J. Cell. Biochem. *39*, 429-41.

Whittaker, S.R., Walton, M.I., Garrett, M.D. & Workman, P. (2004) The cyclin-dependent kinase inhibitor CYC202 (R-roscovitine) inhibits retinoblastoma protein phosphorylation, causes loss of cyclin D1, and activates the mitogen-activated protein kinase pathway. Cancer Res. *64*, 262-272.

Woelfel, T., Hauer, M., Schneider, J., Serrano, M., Woelfel, C., Klehmann-Hieb, E., De Plaen, E., Hankeln, T., Meyer zum Büschenfelde, K.H. & Beach, D. (1995) A p16INK4a-insensitive CDK4 mutant targeted by cytolytic T lymphocytes in a human melanoma. Science *269*, 1281-4.

Woods, K.A., Camacho-Hübner, C., Savage, M.O. & Clark, A.J. (1996) Intrauterine growth retardation and postnatal growth failure associated with deletion of the insulin-like growth factor I gene. N. Engl. J. Med. *335*, 1363-7.

Wu, Y., Cui, K., Miyoshi, K., Hennighausen, L., Green, J.E., Setser, J., LeRoith, D. & Yakar, S. (2003) Reduced circulating insulin-like growth factor I levels delay the onset of chemically and genetically induced mammary tumors. Cancer Res. *63*, 4384-8.

Wüthrich, K. (1986) NMR of proteins and nucleic acids, Wiley, New York

Yamamoto, A., Guacci, V. & Koshland, D. (1996) Pds1p, an inhibitor of anaphase in budding yeast, plays a critical role in the APC and checkpoint pathway(s). J. Cell. Biol. *133*, 99-110.

Yamasaki, H., Prager, D. & Melmed, S. (1993) Structure-function of the human insulin-like growth factor-I receptor: a discordance of somatotrophin internalization and signaling. Mol. Endocrinol. *7*, 681-5.

Yang, E., Zhang, K., Li Cheng, Y. & Mack, P. (1998) Butein, a Specific Protein Tyrosine Kinase Inhibitor. Biochem. Biophys. Res. Comm. *245*, 435-438.

Yang, M., Li, B., Liu, C.J., Tomchick, D.R., Machius, M., Rizo, J., Yu, H. & Luo, X. (2008) Insights into mad2 regulation in the spindle checkpoint revealed by the crystal structure of the symmetric mad2 dimer. PLoS Biol. *6*, e50.

Yu, H. & Rohan, T. (2000) Role of the insulin-like growth factor family in cancer development and progression. J. Natl. Cancer Inst. *92*, 1472-89.

Yu, S. M., Cheng, Z. J. & Kuo, S.C. (1995) Endothelium-dependent relaxation of rat aorta by butein, a novel cyclic AMP-specific phosphodiesterase inhibitor. Eur. J. Pharmacol. *280*, 69-77

Zachariae, W. & Nasmyth, K. (1999) Whose end is destruction: cell division and the anaphase-promoting complex. Genes. Dev. *13*, 2039-58.

Zeng, Y., Forbes, K.C., Wu, Z., Moreno, S., Piwnica-Worms, H. & Enoch, T. (1998) Replication checkpoint requires phosphorylation of the phosphatase Cdc25 by Cds1 or Chk1. Nature *395*, 507–510.

Zhang, L., Zhou, W., Velculescu, V.E., Kern, S.E., Hruban, R.H., Hamilton, S.R., Vogelstein, B. & Kinzler, K.W. (1997). Gene expression profiles in normal and cancer cells. Science *276*, 1268–1272.

Abbreviations and symbols

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
Å	Ångstrøm (10^{-10} m)
aa	amino acid
ALS	amyotrophic lateral sclerosis
APC/C	anaphase-promoting complex or cyclosome
APS	ammonium peroxodisulfate
ATP	adenosine-5'-triphosphate
BMMH	Buffered Minimal Methanol Histidine medium
bp	base pair
BSA	bovine serum albumin
CAP	cyclase associated protein
CAPS	N-cyclohexyl-3-aminopropanesulfonic acid
Cdc	Cell division cycle protein
CDK	cyclin-dependent kinase
cDNA	complimentary DNA
CIN	chromosomal instability
CIP	calf intestinal alkaline phosphatase
Cip/Kip	CDK interacting protein and kinase inhibitory protein family
CKI	CDK inhibitor
CR	a cysteine-rich region
Da	Dalton (g mol-1)
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E. coli	*Esherichia coli*
EDTA	ethylenediamine tetraacetic acid
FnIII	fibronectin type III domains
g	gravity (9.81 m s-2)
GH	growth hormone

GSH	reduced glutathione
GSSG	oxidized glutathione
HDAC	histone deacetylase
HEPES	N-(2-hydroxyethyl)piperazine-N`-(2-ethanesulfonic acid)
HSQC	heteronuclear single quantum coherence
Hz	Hertz
ID	insert domain
IGF	insulin-like growth factor
IGFBP	IGF binding protein
IGF-1R	IGF receptor type I
IGF-2R	IGF receptor type II
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl-β-thiogalactopyranoside
IR	insulin receptor
IRR	IR-related receptors
IRS	insulin receptor substrate(s)
ITC	isothermal titration calorimetry
K_D	dissocation constant
L1	leucine-rich repeat domain
L2	second leucine-rich repeat domain
LB	Luria-Broth medium
Mad1	mitotic arrest deficient 1 protein
Mad2	mitotic arrest deficient 2 protein
MAPK	mitogen-activated protein kinase
MIN	microsatellite instability
MM	minimal medium
MW	molecular weight
NiNTA	nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
OD	optical density
P. pastoris	*Pichia pastoris*
P3K	phosphatidylinositol 3-kinase
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline

PEG	polyethylene glycol
pI	iso-electric point
PI-3	phosphoinositol-3 kinase
Plk1	polo like kinase 1
PP2A	protein phosphatase 2A
ppm	parts per million
pRb	retinoblastoma protein
RNA	ribonucleic acid
SAC	spindle assembly checkpoint
SB	sample buffer
SDS	sodium dodecyl sulfate
Sgo	shugoshin
Smc	structural maintenance of chromosomes subunits
SMS	somatostatin
SUMO	Small Ubiquitin-like Modifier
TAE	Tris-acetate-EDTA buffer
TCA	trichloroacetic acid
TEMED	N,N,N',N'-tetramethylethylendiamine
TEV	Tobacco Etch Virus
TGF-β	transforming growth factor-β
TNF	Tumor Necrosis Factor
Tris	tris(hydroxymethyl)aminomethane
UCN-01	7- hydroxystaurosporine
YNB	yeast nitrogen base
YP	yeast extract peptone

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