Study of Proteins as Drug Targets by NMR spectroscopy

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim
Fachbereich Biochemie, Chemie und Pharmazie der
Johann Wolfgang Goethe - Universität
zu Frankfurt am Main

von

Sridhar Sreeramulu

aus
Ambur (Indien)

Frankfurt 2009
D30

vom Fachbereich Biochemie, Chemie und Pharmazie der
Johann Wolfgang Goethe-Universität als Dissertation angenommen.

DEKAN: Prof. Dr. Dieter Steinhilber

1. GUTACHTER: Prof. Dr. Harald Schwalbe
2. GUTACHTER: Prof. Dr. Volker Dötsch

DATUM DER DISPUTATION: 2009

To my parents and wife

This thesis was prepared under the supervision of Prof. Dr. Harald Schwalbe between Sept 2003 and May 2009 at the Institute for Organic Chemistry and Chemical Biology of the Johann-Wolfgang Goethe-University Frankfurt.

Contents

List of Figures

List of Tables

1 Overview and Summary

The following thesis is concerned with the study of proteins and protein-protein complexes of biomedical importance by Nuclear Magnetic Resonance spectroscopy (NMR). Understanding how biological systems operate from the level of a single protein to more complex protein-protein complexes, and finally to the transduction of signal holds the key for finding cure for many diseases. Cancer is one such disease which is characterized by unregulated cell division often promoted by protein kinases [Teicher, 2000]. Therefore, protein kinases are often targeted by small molecule inhibitors, albeit with the problem of insufficient target selectivity due to the high sequence similarity between different kinase family members [Fabbro et al., 2002; Noble et al., 2004].

Akt/PKB is a protein kinase that has been found frequently in a constitutively active form in human cancers. Akt/PKB is a clinically validated target and hence small molecule inhibitor towards this kinase are highly desirable [Brognard et al., 2001; Hanada et al., 2004].

Akt/PKB cannot easily be expressed in *E. coli*, and hence cAMP dependent kinase (PKA) which has high sequence homology to Akt/PKB, has been used as a surrogate kinase for drug design.

As alternative route, drugs that inhibit the molecular chaperone Hsp90 are currently in clinical trials due to their ability to promote degradation of many kinases [Whitesell and Lindquist, 2005]. More recently Cdc37, a cochaperone of Hsp90 in mammalian cells, has been identified to target protein kinases. It is upregulated in various cancers [Lee et al., 2002; Pearl, 2005]. The protein-protein (Hsp90-Cdc37)

complex forms with a $K_D = 1.2\ \mu M$, and is considered to mediate carcinogenesis by stabilizing a variety of different oncogenic kinases in malignant cells.

The method in this thesis to study these proteins or protein-protein complex in high resolution is a combination of liquid-state NMR spectroscopy and X-ray crystallography. It is important to characterize this interaction at very high resolution as it will enable us in identifying 'Hot spots' in the protein-protein interaction interface and in further developing small molecule inhibitors [Wells and McClendon, 2007].

The first part of the thesis introduces the reader to the world of Structural proteomics and its importance. Chapter 2 discusses various projects undertaken by the world community in the post human genome era, in order to obtain the three-dimensional structures of a substantial set of proteins of an organism, the 'proteome'. Chapter 2 also gives special emphasis on Structural Proteomics in Europe (SPINE), which decided to limit its study to proteins and protein-protein complexes of biomedical importance and also explicitly provides PDB statistics about three-dimensional structures of high molecular weight proteins solved by NMR spectroscopy. The importance of the study of Cancer-related proteins like protein kinases, Hsp90 and Cdc37, are introduced in Chapter 3. Chapter 4 gives the reader a taste and importance of developing small molecule inhibitors in disrupting protein-protein interactions in order to prevent diseases. Finally, Chapter 5 reviews various NMR spectroscopy methods used to study protein-protein interaction. This chapter also introduces the reader the importance of protein-protein interaction in the cellular processes.

The second part of the thesis, which is a cumulative part, encompasses the original research work conducted. Chapter 6 describes how NMR spectroscopy was used to access the correct folding of protein kinase A (PKA) and the mutants. Further, it reports how mutations of phosphorylation sites in PKA affect the expression level, stability and activity of the kinase.

Achievements:

- Establishment of PKA with three phospho-mutations, can easily be expressed in *E. coli*, which otherwise is a major obstacle for most of the kinases
- Establishment of the correct fold of the three phospho-mutants of the 40 kDa PKA by NMR

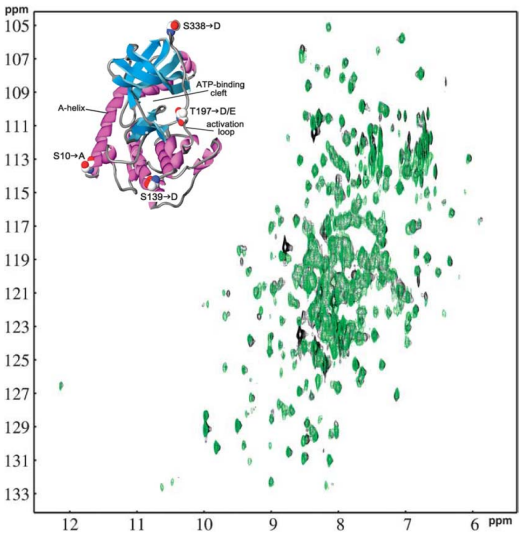


Figure 1.1: Folding and activity of cAMP-dependent protein kinase mutants. Comparison of TROSY spectra of wild-type PKA (black) with "dephospho"-PKA mutant, PKA-3P(green) shows that the mutants are correctly folded. For more details see Chapter 6.

- Revelation that there will be no mixtures of isoforms due to differential phosphorylation, since there is now only one phosphorylation site in the activation loop, which makes it highly suitable for studying protein-ligand interactions
- Establishment of the fact that these mutants are active and well folded which can now be highly advantageous for further functional and structural studies

Chapter 7 exemplifies how we achieved the structure determination of a 40 kDa human Cdc37$_M$-Hsp90$_N$ protein-protein complex by using NMR spectroscopy in combination with X-ray crystallography. Further, by using NMR spectroscopy we could pinpoint the 'Hot Spot' in the large interaction interface of this protein-protein complex.

Achievements:

- Establishment of the domain boundaries, purification and NMR conditions for the, rather challenging, 45 kDa protein Cdc37
- Achievement of complete (99.5 %) backbone resonance assignment of the Hsp90 binding domain of Cdc37
- Determination of the structure of the human Hsp90 binding domain of Cdc37 by X-ray crystallography in collaboration with Prof. Dr. Roy Lancaster
- Mapping of the interaction interface of the Cdc37-Hsp90, by using multitude of NMR techniques, which sets an example for further studies of large protein-protein complexes

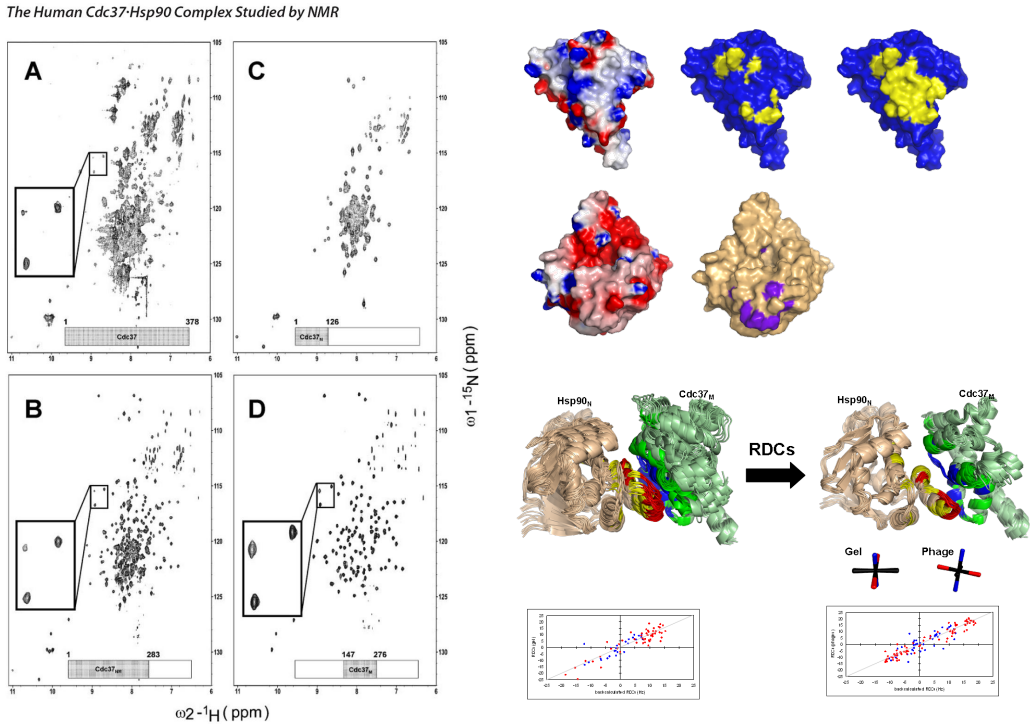


Figure 1.2: NMR studies of human Cdc37$_M$-Hsp90$_N$ complex. A combination of domain optimization, NMR spectroscopy methods (resonance assignment, chemical shift perturbation mapping, cross-saturation transfer, RDC's), X-ray crystallography and docking using HADDOCK was used to determine the solution structure of 39 kDa human Cdc37$_M$-Hsp90$_N$ complex. For more details see Chapter 7.

- Successful usage of residual dipolar couplings (RDCs) for more accurate structure determination of the protein-protein complex
- Determination of the structure of 40 kDa human Cdc37-Hsp90, by using a combination of 'state-of-the-art' techniques NMR, X-ray and docking
- PDB statistics show that till date there are only 49 structures above 39 kDa out of 7728 structures solved by NMR. The structure of Cdc37-Hsp90 complex is one of the 49 structures solved by NMR
- Identification of the 'Hot Spot' in the large interaction interface by NMR and also confirmed it further, by a combination of single point mutation and NMR

Finally, Chapter 8 describes the achievement by NMR that celastrol, a recently identified triterpene, targeting Hsp90, in fact binds to Cdc37 and disrupts the Cdc37-Hsp90 complex. Further, it also describes the molecular mechanism of inhibition of Cdc37 as solved by NMR.

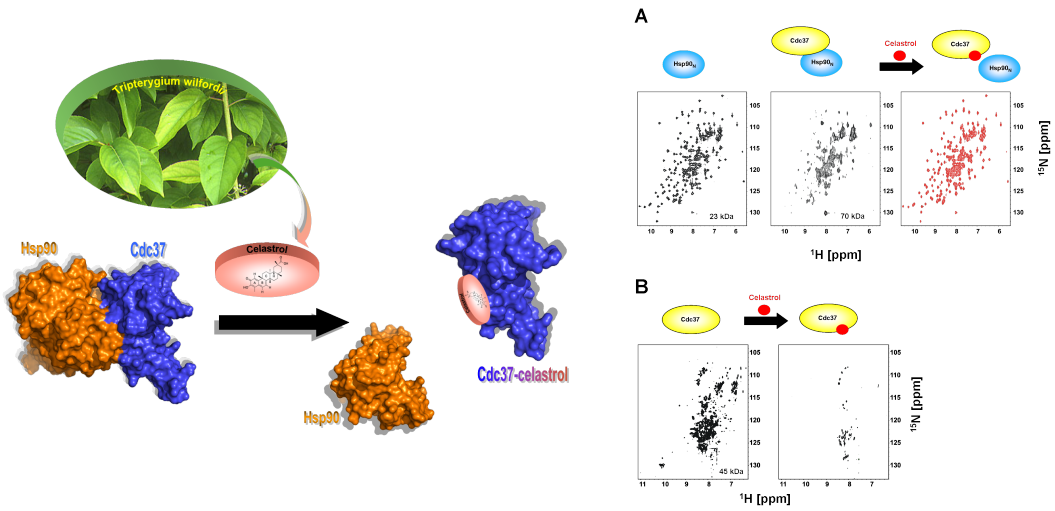


Figure 1.3: Celastrol binds to Cdc37. The cell division cycle protein 37 (Cdc37) and the heat shock protein (Hsp90) are molecular chaperones, crucial for the folding and stabilization of protein kinases including the oncogenic kinases. Here we show by NMR that celastrol, a recently identified triterpene targeting Hsp90, in fact binds to Cdc37 and disrupts the Cdc37-Hsp90 complex. Celastrol inactivates Cdc37 through a thiol-mediated mechanism. For more details see Chapter 8.

Achievements:

- Discovery of Cdc37 as the right target for celastrol within the protein-protein (Cdc37-Hsp90) complex by NMR
- Establishment of the exact mechanism of interaction between celastrol and Cdc37 by NMR
- Revelation that the N-terminal and the Middle domain of Cdc37 are responsible for the binding of celastrol
- Importantly, this study will shift the focus of the medicinal chemistry efforts towards the correct target within the protein-protein complex.

LIST OF PUBLICATIONS:

1. Folding and activity of cAMP-dependent protein kinase mutants: Thomas Langer, Sridhar Sreeramulu, Martin Vogtherr, Bettina Elshorst, Marco Betz, Ulrich Schieborr, Krishna Saxena, Harald Schwalbe, *FEBS Letters*, 2005, **579**, 4049-4054.

2. ^{1}H, ^{13}C and ^{15}N backbone resonance assignment of the Hsp90 binding domain of human Cdc37: Sridhar Sreeramulu, Jitendra Kumar, Christian Richter, Martin Vogtherr, Krishna Saxena, Thomas Langer, Harald Schwalbe, *Journal of Biomolecular NMR*, 2005, **32**, 262.

3. The Human Cdc37.Hsp90 Complex Studied by Heteronuclear NMR Spectroscopy: Sridhar Sreeramulu, Hendrik Jonker, Thomas Langer, Christian Richter, Roy Lancaster, Harald Schwalbe, *Journal of Biological Chemistry*, 2009, **284**, 3885-3896.

4. Molecular Mechanism of Inhibition of the Human Protein complex Hsp90-Cdc37, a Kinome Chaperone-Cochaperone, by Triterpene Celastrol, Sridhar Sreeramulu, Santosh Lakshmi Gande, Michael Göbel, Harald Schwalbe, 2009, *Angewandte Chemie International Edition English*, 2009, in press.

2 Structural Proteomics: An Overview

2.1 Introduction

Successful completion of the Human genome sequencing has opened up new avenues in the field of biology and medical sciences [Lander et al., 2001; Venter et al., 2001]. This achievement was considered and foreseen to revolutionize and reveal new treatment strategies for diseases and broaden our understanding of Biology. Forgotten however, has been the fact that these meager complete sequences (three billion bases), does not *per se* mean anything, unless they are given a meaning, by figuring out the proteins that these genes encode and what they do for a living. Further, understanding how these proteins collaborate (protein-protein interaction) with each other, in order to carry out the complex cellular events, is the real challenge ahead. Hence, we move from the completion of 'Genome' towards the task of completing the 'Proteome (Structural genomics/Structural proteomics)' [Fields, 2001].

Unlike genomics, proteomics comes with much more bigger challenge, due to the fact that proteins are much more versatile than nucleic acids. A single gene can code for multiple proteins and further these proteins can undergo modifications (such as phosphorylation, glycosylation, acetylation, ubiquitation) or a single protein can have different functions under different physiological conditions or multiple proteins could also do the same function. All these arguments imply that the human proteome will be an order of magnitude more complex than the genome. Hence, the task ahead needs a confluence of scientists from various disciplines, from geneticists, biochemist to structural biologists. The last two decades, the pre-proteomic era, have witnessed characterization of thousands of proteins, which is

a result of the confluence of scientists from various disciplines, advancements in technology to study these proteins involved in cellular processes and also a result of conservation of fundamental mechanisms between the organisms. In comparison to traditional biological research which was mainly based on specific models of cellular behavior, proteomics aims at more systematic studies.

In the post human genome era it was envisaged that, in analogy to sequencing (genomics), structural genomics is defined as the determination of the three dimensional structures of all the proteins coded by a genome. The successful use of high-throughput sequencing methods now encouraged to use the same technology to solve the three dimensional structures of proteins. But solving the structures of all proteins is an impractical goal because many of the proteins are not still amenable to purification, in a form suitable for structure determination or in fact are inherently unstructured. Hence different organizations across the world defined their structural genomics goals and started a focused research in achieving those targets.

2.2 Global structural efforts and target selection strategies

In the mid-to-late 1990s in both Japan and USA, the concept of structural genomics was already taking birth, encouraged by the success of the use of High-throughput (HTP) sequencing methods for sequencing the genome. It was envisaged that similar application of HTP could even be used to achieve the three-dimensional structures of substantial fraction of proteins of a given organism [Kobe et al., 2008; Sussman and Silman, 2003].

2.2.1 Japan

In October, 1998, the Protein Research Group was established at the RIKEN Genomic Sciences Center in Japan. Since its establishment it has elucidated the structures and functions of many important proteins. In 2002, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) started this national project to obtain deep insight into the biological network by solving over 3,000 protein structures (“National Project on Protein Structural and Functional Analyses” (the “Protein 3000” Project, 2002-2006) and determining functions of biological and medical importance. While almost 90 laboratories in universities and research institutions

joined this project, the RIKEN Structural Genomics/Proteomics Initiative (RSGI) conducted systematic and comprehensive research to elucidate protein structures and functions at the RIKEN Genomic Sciences Center and the SPring-8 Center. RSGI chose primary protein targets which were involved in signal transduction, nucleic acid binding and were of medical importance from various species like humans, mice, and *Arabidopsis thaliana*. In March 2007, " Project 3000" was declared over and aim achieved with RSGI solving about 2500 protein structures and 500 structures coming from other research centers across Japan. Currently, Systems and Structural Biology Center (SSBC) established in April, 2008, is moving forward with all the experiences gained from the previous projects in order to solve the structures of more complex proteins. RIKEN through Systems and Structural Biology Center (SSBC), aims to build a bridge between life science and medical science, to expand the logical design of biomolecular mechanisms and to increase predictability in life science.

2.2.2 United States of America (USA)

On April 24, 1998, a one day meeting (as a follow up of larger meeting held in the same year at Argonne National Laboratory) was held to discuss issues related to a genome-directed Protein Structure Initiative (PSI) and finally in 2000, the Protein Structure Initiative (PSI) was started and funded together by the National Institutes of Health (NIH) and National Institute of General Medical Sciences (NIGMS). The main aim of this structural genomics initiative was the structure determination of a representative from every possible protein fold or motif. The total number of different folds varies between 1000 to nearly 10,000, spread across various species; hence the structures of these projects would come from different genomes. The structure with respect to one protein fold from one organism would then be used to model related structures in other organisms by homology modeling. To discover all protein folding motifs, it is estimated that on the order of 3,000-5,000 new protein structures must be determined experimentally. PSI is a 10 year project which will be implemented at different phases. PSI phase I (PSI-1 or also known as Pilot phase) was initiated in year 2000 and 2001, were conducted in nine centers, and were completed in 2005, with more than 1300 structures having been deposited in the PDB. Among this more than 700 structures had less than 30 percent sequence identity. All-in-all PSI-1 achieved its purpose of developing new tools, technologies and methodology to increase the success rates and lower the costs of structure determination. Phase II (PSI-2) studies were immediately started with a goal of producing 4000 new structures within five years. Within two years of operation

PSI-2 deposited 1200 structures in the PDB, thus surpassing five-year combined production output of the PSI-1. Currently, PSI-2 is still continuing with the aim basically the same as PSI-1, but with more focused and clearly defined objectives in order to attain its overall goal.

2.2.3 North America

In 2003, a major joint venture between government and industry resulted in the establishment of the Structural Genomics Consortium, an international project funded by governments of Canada and Sweden, the Wellcome Trust in the UK and industries, with laboratories in Oxford, Stockholm and Toronto. The Structural Genomics Consortium (SGC) is a not-for-profit organization that aims to determine the three dimensional structures of proteins of medical relevance, and place them in the public domain without restriction. It focused on proteins and protein families from human and apicomplexan (e.g. *Plasmodium falciparum* which causes malaria) that are either potential drug targets or have been implicated in human disease processes. In the period from July 2004 to June 2007, the SGC deposited the structures of 450 proteins from its Target List of ∼2,000 proteins, to achieve a deposition rate of 200 structures per year. SGC's areas of interest include proteins of signaling pathways, protein kinases and human cytosolic sulfur transferases. The consortium has solved more than 80 novel human kinase structures till date. SGC has also studied a total of 1008 genes from *P. falciparum* and has resulted in the determination of more than 36 structures. The results are now being actively applied for the development of vaccines and small molecule therapeutics. Currently, it has deposited 716 structures out of 4594 targets, and the rest are in the different stages of structure determination.

2.2.4 Structural proteomics in Europe (SPINE)

On 1st October, 2002, SPINE was started as a three-year project funded by the European Union Fifth Framework Program (EU FP5) [Stuart et al., 2006]. Unlike all the previous structural genomics programs, SPINE was named differently as Structural Proteomics program in order to differentiate itself from others. The selection of targets was driven by the motto "human health targets". It aimed mainly at solving proteins and protein complex structures directly relevant to human health and diseases. SPINE aimed at developing technologies for HTP structure determination both by X-ray and NMR involving proteins and protein complexes from eukaryotes and prokaryotes. Apart from trying to establish

European centers of Excellence it also aimed at training young scientists and technicians to take part in the above program. SPINE chose its targets mainly in two phase, first involving identification of targets with possible biomedical importance, and the second being an assessment of whether the target is amenable for structural studies. SPINE produced 375 structures of which 308 were unique proteins and the rest were protein-ligand complexes. It should not be forgotten that one of the challenging aspects of SPINE project was to focus on Human and other eukaryotic proteins that are potentially of high biomedical importance. A significant high number of 170 structures solved by SPINE come from eukaryotic origin. Current version of SPINE is SPINE2-complexes funded by the European Union Framework 6. The full title of the project is "From Receptor to Gene: Structures of complexes from signalling pathways linking immunology, neurobiology and cancer". It aims to focus mainly on protein-protein and protein-nucleic acid complexes of biomedical importance. Frankfurt has been a member of SPINE and SPINEII as subcontractor. In this thesis, we focus our study on proteins related to cancer viz., protein kinases (protein kinase A (PKA)) and human Hsp90-Cdc37, a kinome chaperone-cochaperone complex.

In summary, whatever was the individual aim of different Structural genomics/Structural proteomics (SG/SP) program throughout the world, they have made and continue to make a major contribution to the development of technology for all aspects of structural biology ranging from identification of target genes to refinement of structures using X-ray and NMR. Table 2.1 summarizes the list of Major Structural Genomics Organizations along with their target selection criteria. Table 2.2 summarizes the overall achievement by the SG/SP collectively throughout the world.

Table 2.1: List of major structural Genomics programs

	Consortium	Target Selection Criteria
1	Protein Structure Initiative, USA	Structure determination of a representative of every possible fold
2	Project 3000, Japan	Novel sequences and biologically important or human health related
3	Structural Genomics Centers, Canada, Sweden, UK	Human proteins related to diseases and human pathogens
4	Structural Proteomics in Europe (SPINE), UK	Human health related proteins (cancer-related proteins, Immune defense, neuronal development and neurodegenerative diseases)

2.3 Role of NMR in structural proteomics

Nuclear Magnetic Resonance (NMR) is a powerful tool to determine structures of proteins in solution at atomic resolution. It neatly complements X-ray crystallography, as many of the proteins do not produce diffraction quality crystals and thus they are solely amenable to be determined by NMR [Sussman and Silman, 2003]. Noteworthy is the fact that NMR can be equally used for both eukaryotic and prokaryotic proteins, while eukaryotic proteins crystallize less frequently than prokaryotic proteins. As of March 6, 2009, there are about 1748 structures solved by NMR within the Structural proteomics program throughout the world, as compared to 6477 total structures solved. This attests the fact that nearly 27 % (Figure 2.1) of the total structures solved comes from NMR. The success and importance of NMR as a structure determination tool, is also due to the fact that there is no correlation between the 'crystallisability' and the quality of its 2D ^{1}H,^{15}N HSQC NMR spectrum, which shows one signal for each amino acid residue. Hence many proteins giving rise to a high quality NMR spectra do not produce diffraction quality crystals and vice versa.

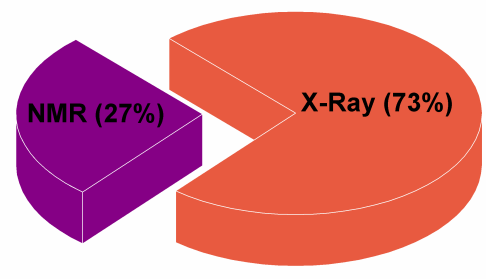


Figure 2.1: Structural proteomics contribution by X-ray and NMR. Protein structure production by X-ray or NMR for all Structural Proteomics Centers in the world (March 2009). 1748 structures has been solved by NMR from a total of 6477 solved. Data obtained from http://targetdb.pdb.org/statistics/TargetStatistics.html (March2009) also see Table 2.2.

Currently, HTP high-quality NMR structural determination is limited to proteins of molecular weights less than 25 kDa. This is because NMR lines increasingly broaden when the overall tumbling of a protein slows down with increasing molecular weight, while the number of resonance lines concomitantly increases with size. HTP NMR structure determination involves preparation of $^{13}C/^{15}N$-labeled proteins, rapid acquisition of NMR data, (semi-) automated data analysis and structure calculation and finally structure validation. The expression of labeled proteins for HTP is now being achieved either by heterologous expression in *E. coli* cells [Acton et al., 2005; Lundstrom, 2007; Yee et al., 2006], wheat-germ based cell-free expression [Kigawa et al., 2004; Tabb et al., 2002; Vinarov et al., 2006] and also recently developed cell-free protein expression using *E. coli* cell extract [Kigawa et al., 2004]. Rapid NMR data acquisition is mainly aimed at "sensitivity limited

Status	Total number of Targets	(%) Relative to "Cloned" Targets	(%) Relative to "Expressed" Targets	(%) Relative to "Purified" Targets	(%) Relative to "Crystallized" Targets
Cloned	140040	100.0	-	-	-
Expressed	93803	67.0	100	-	-
Soluble	36428	26.0	38.8	-	-
Purified	33463	23.9	35.7	100.0	-
Crystallized	11828	8.4	12.6	35.3	100.0
Diffraction-quality crystals	5921	4.2	6.3	17.7	50.1
Diffraction	5326	3.8	5.7	15.9	45.0
NMR assigned	1803	1.3	1.9	5.4	-
HSQC	3517	2.5	3.7	10.5	-
Crystal structure	4271	3.0	4.6	12.8	36.1
NMR structure	1721	1.2	1.8	5.1	-
In PDB	6276	4.5	6.7	18.8	39
Work stopped	35183	-	-	-	-
Test target	61	-	-	-	-
Other	8612	-	-	-	-

Table 2.2: Success rates for all structural proteomics centers (March 2009)

data acquisition regime" in which the instrument time is invested only to the extent that the signal-to-noise ratios of most peaks reach reliable peak identification. The techniques like G-matrix Fourier Transform (GFT) NMR spectroscopy [Kim and Szyperski, 2003], simultaneously acquired $^{13}C/^{15}N$ resolved NOESY [Shen et al., 2005], L-optimization (Longitudinal relaxation optimization) have now significantly reduced the time of data acquisition [Pervushin et al., 2002]. (Semi-) Automated data analysis and structure determination for HTP structural determination uses program like AUTOASSIGN [Huang et al., 2005], CYANA [Güntert, 2003] and AUTOSTRUCURE [Huang et al., 2006].

2.4 Protein Data Bank (PDB) statistics for structures solved by NMR

Recent technological advances in NMR have made it possible to contribute significantly to the structure determination of proteins. One of the major limitations of structure determination by NMR is the molecular weight of proteins. Here we summarize (Figure 2.2, Table 2.3) the current status of the protein data bank with respect to structures solved by NMR, correlating it with the molecular weight.

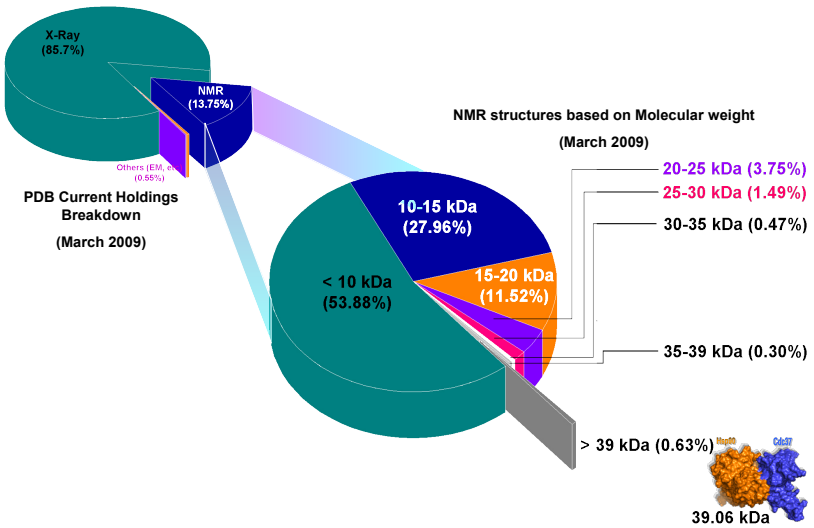


Figure 2.2: PDB statistics. The above picture summarizes the number of structures contributed by different experimental techniques. It also gives a detailed picture of the correlation between molecular weight and number of structures solved by NMR. In this thesis, we study the protein-protein complex (Hsp90-Cdc37) which has a molecular weight of 39 kDa. Data obtained from `http://www.rcsb.org/pdb/static.do?p=general_information/pdb_statistics/index.html` (March2009) also see Table 2.3.

		Molecule type				
		Proteins	**Nucleic acids (NA)**	**Protein-NA complexes**	**Other**	**Total**
Experiment method	**X-Ray**	44950	1123	2064	24	48161
	NMR	6742	837	142	7	7728
	Electron Microscopy	146	16	59	0	221
	Others	96	5	4	2	107
	Total	51934	1981	2269	33	56217

Table 2.3: PDB current holdings breakdown (March 2009)

3 Importance of Study of Various Protein Families of Biomedical Relevance

3.1 Introduction

Proteins perform multiple crucially important functions. They are involved in almost every aspect of biological activity and process, ranging from providing mechanical support viz., actin and myosin from contractile machinery to more complex cellular process viz., signal transduction pathways involving kinases, phosphatases, surface receptors.

The human genome is estimated to encode 30,000 to 40,000 proteins from a relatively low number of genes encoding proteins [Venter et al., 2001]. Only 1.1 to 1.5 percent of the human genome encodes proteins. Computational analysis of the predicted protein-coding sequences shows that over 40 percent of the proteins cannot be ascribed a molecular function by methods that assign proteins to known families and hence termed as 'Unknown function'. Identification of these unknown families will be a major focus for many laboratories. Major classes of proteins encoded by human genome are graphically represented in Figure 3.1.

Advances in molecular genetics reveal that human diseases are mostly dictated by the changes in one or more of these proteins. Hence, an important goal of molecular medicine is to identify such proteins whose presence, absence, or deficiency is associated with specific physiologic state or disease. A classic example is cystic fibrosis, an inherited condition which involves the protein cystic fibrosis transmembrane conductance regulator (CFTR). In 65 % of reported cases a deletion of single amino acid (phenylalanine) at position 508 in the 1500 residue protein

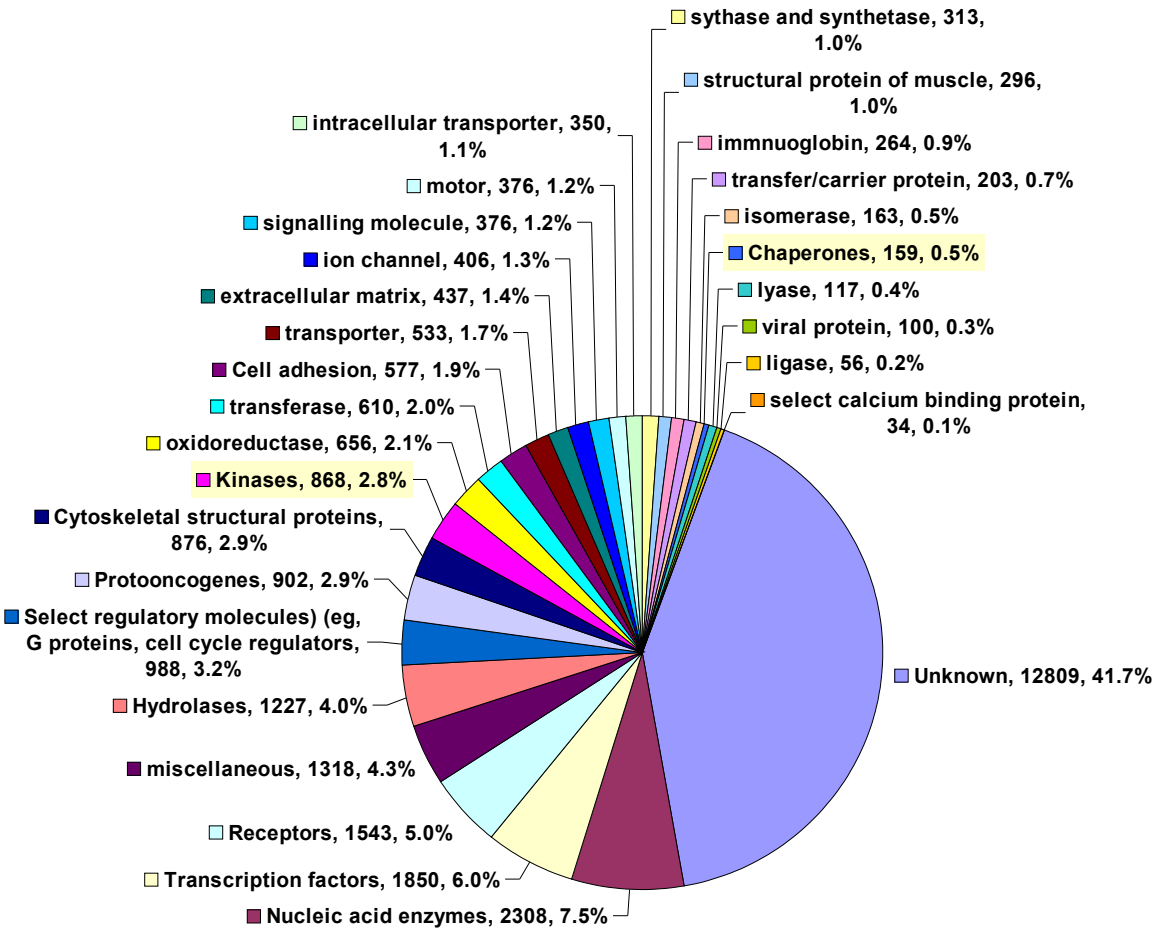


Figure 3.1: Molecular functional distribution of human genome. Distribution of the molecular functions of human genes. Each slice lists the number and percentages of human gene functions assigned to a given category of molecular function [Venter et al., 2001].

results in the disease. A change from glutamic acid to valine of the sixth amino acid out of 574 amino acids in haemoglobin results in the disease sickle cell anemia. Study of proteins and also protein-protein interaction during the signal transduction holds the key for understanding the etiology of many diseases viz., cancer (protein kinases, phosphatases, chaperones etc).

3.2 Cancer

One of the major challenges of the 21st century is to overcome the challenges posed by the disease cancer [Lewis, 2006]. Cancer is uncontrolled proliferation of cells that have lost their normal regulated cell division often in response to genetic or environmental trigger. The development of cancer is a multistep and

multifactorial phenomenon, often requiring several years. The accumulation of the basic knowledge from the past few decades has resulted in significant progress in understanding the molecular basis of cancer. Although the process of tumorigenesis is incompletely understood, it is clear that successive accumulation of mutations in key genes is the force that drives towards cancer. These mutations can occur as either,

- a multistep step process due to aging and such tumor can be considered as Darwinian evolution on a microscopic scale with each successive step overcoming the social rules of a normal cell and hence called "clonal evolution" [Nowell, 1976]
- exposure to carcinogens either physical or chemical agents [Hennings et al., 1993]
- epigenetic changes viz., hypermethylation of adenomatous polyposis coli (APC) gene in colon cancer [Slattery et al., 2001]

3.2.1 Molecular basis of cancer phenotypes [Abraham et al., 2003]

Differentiating the tumor cells from the normal cells is the key to understanding the cancer cells and ultimately leads to therapies that can target tumor cells. Some of the features are,

- **Immortality:** Cancer cells proliferate indefinitely due to abnormal up-regulation of telomerase activity. Hence, telomerase is a useful therapeutic target. A second cause could arise from mutations in the Retinoblastoma (Rb) [Zheng and Lee, 2001] and p53 genes which are responsible for tumor suppressor activity.
- **Decreased dependence on the growth factors to support proliferation:** In contrast to the normal cells which obtain their growth factors extracellularly, the tumor cells produce their own growth factors. These growth factors then bind to, and continuously stimulate the receptors present on the same tumor cells, which results in autocrine stimulation (self-generated proliferative signal) that drives proliferation. Epidermal growth factor receptor (EGFR) plays a major role in the progression of most human epithelial tumors [Salomon et al., 1995].
- **Loss of anchorage-dependent growth and altered cell adhesion:** Tumor cells in contrast to normal cells can grow even in suspensions or on semi-solid

agar gels as they have lost the anchorage-dependent growth. This results in migration of the tumor cells to invade and metastasize foreign tissues. The detachment of the tumor cells from its parent tissue is driven by the enzymes Matrix metalloproteinases (MMPs), which otherwise are absent in normal adult cells. Hence, MMPs have become targets for developing inhibitors.

- **Cell cycle and loss of cell cycle control:** The cell cycle can be considered as the heart of the proliferation. All events must be completed in a timely and sequence specific fashion in order to function as a normal cell. It is a complex process which involves many regulatory steps and signal transduction pathways, which also harbor oncogenes and tumor suppressor genes. Cell division is divided into four phases G1, S, G2 and M. Out of these, G1 is the only nonreplicative phase or in other terms the resting phase and needs strong presence of growth factors and nutrients in order to exit to the next phase. The phase transition from G1 to S is tightly regulated, and is misregulated in neoplastic cells resulting in uncontrolled proliferation. *In this thesis, we study one of the proteins (Cdc37) which is required for the passage of the G1 phase of the cell cycle and is implicated as target for developing therapeutics [Reed, 1980].*

- **Apoptosis and reduced sensitivity to apoptosis:** Apoptosis is a genetically controlled form of cell death required to maintain the correct number of cells in adult life. Any imbalance in this process will also lead to uncontrolled proliferation of cells. Hence, identifying and understanding apoptotic process will give new targets for therapeutic manipulation. Examples include Tumor Necrosis Factor Receptor 1 (TNFR1) and several other downstream proteins involved in the signal transduction pathway [Chen and Goeddel, 2002]. Bcl-2 (B-cell lymphoma-2), which is overexpressed in cancer cells and has become an important target for developing drugs [Zamzami et al., 1998].

- **Increased genetic instability:** The gain or loss of one or more specific chromosomes results in genetic instability, which is a hallmark of cancer cells. Acquisition of extra chromosome is one mechanism by which extra copies of a growth promoting gene can be acquired by the tumor cells, providing them the selective growth advantage. One example is Philadelphia chromosome, which results from the translocation of chromosome 9 and 22. Two genes (c-Abl and Bcr, a tyrosine kinase and a GTPase activating protein (GAP)) of completely unrelated proteins are spliced together, forming a chimeric protein that results in powerful and constitutively active kinase that drives proliferation [Stam et al., 1985].

- **Angiogenesis:** Lack of blood vessels will restrict the growth of tumor cells to only a few millimeters. The hypoxic environment created inside the tumor cells upregulates the expression of the transcription factor HIF1(Hypoxia Inducible Factor) which is kept at a very low concentrations by the presence of VHL tumor suppressor protein in normal cells. HIF1 along with Vascular Endothelial Growth Factor (VEGF), in conjunction with other cytokines, promotes neovascularization of tumors, allowing them to overcome the oxygen diffusion and grow bigger in size. VEGF has become a clinically valid therapeutic drug target in suppressing tumor [Ferrara, 2001; Folkman, 1971].

3.2.2 Oncogenes as therapeutic targets [Croce, 2008]

Table 3.1 contains a summary of the targets and drugs (small molecules and monoclonal antibodies) being used in the treatment of a variety of human cancers.

Table 3.1: Summary of cancer targets and drugs

Anticancer Drug	Target	Disease
Monoclonal antibodies		
Trastuzumab (Herceptin, Genentech)	ERBB2	Breast cancer
Cetuximab (Erbitux, ImClone)	EGFR	Colorectal cancer
Bevacizumab (Avastin, Genentech)	VEGF	Colorectal cancer, non-small-cell lung cancer
Small molecules		
Imatinib (Gleevec, Novartis)	ABL, PDGFR, KIT	Chronic myelogenous leukemia, gastrointestinal stromal tumors, chordoma
Gefitinib (Iressa, AstraZeneca)	EGFR	Non-small-cell lung cancer
Erlotinib (Tarceva, Genentech)	EGFR	Non-small-cell lung cancer
Sorafenib (Nexavar, Bayer/Onyx)	VEGFR, PDGFR, FLT3	Renal cell carcinoma
Sunitinib (Sutent, Pfizer)	VEGFR, PDGFR, FLT3	Gastrointestinal stromal tumors, renal-cell carcinoma

3.3 Protein Kinases

The enzymes carrying out phosphorylation reactions are called protein kinases (Figure 3.2). Protein kinases are vital for many regulatory cellular processes such as cell division, growth and death. These enzymes transfer phosphate from adenosine triphosphate (ATP) onto target proteins, thereby activating or deactivating them.

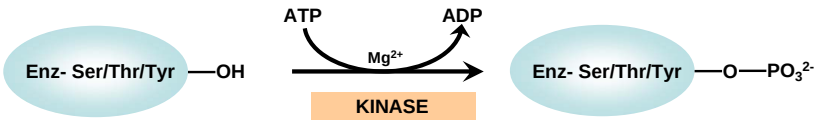


Figure 3.2: Protein kinases. Covalent modification of a regulated enzyme by phosphorylation of either serine, threonine or tyrosine residues.

3.3.1 Classification of the kinome superfamily [Hanks and Hunter, 1995; Sowadski and Epstein, 2000]

There are three major classes of protein kinases with differing protein targets:

1. Serine/Threonine protein kinases transfer a phosphate from ATP to a serine or threonine residue in the target protein and are generally associated with cytoplasmic signaling events. Example: Protein kinase A (PKA)

2. Protein tyrosine kinases transfer phosphate from ATP to tyrosine residues in the target protein and are generally associated with cell surface receptors that become activated after binding a growth factor or other ligand. Example: BCR-ABL kinase

3. Dual specificity protein kinases transfer phosphate from ATP to both threonine and tyrosine residues of the target protein. This group of protein kinases is small in number. Examples: ERK1 and ERK2.

Another classification of protein kinase superfamily is based on the structure and function. The sole consideration in this type of classification was similarity in kinase domain amino acid sequence. Accordingly four groups were designated:

1. the **AGC group**, which includes the cyclic-nucleotide-dependent family (PKA and PKG), the protein kinase C (PKC) family, the β-adrenergic receptor kinase (βARK) family, the ribosomal S6 kinase family and other close relatives.

2. the **CaMK group**, which includes the family of protein kinases regulated by calcium/calmodulin, the Snf1/AMPK family and other close relatives.

3. the **CMGC group**, which includes the family of cyclin-dependent kinases, the Erk (MAP) kinase family, the glycogen synthase 3 (GSK3) family, the casein kinase II family, the Clk (Cdk-like kinase) family and other close relatives.

4. the 'conventional' protein-tyrosine kinase (PTK) group.

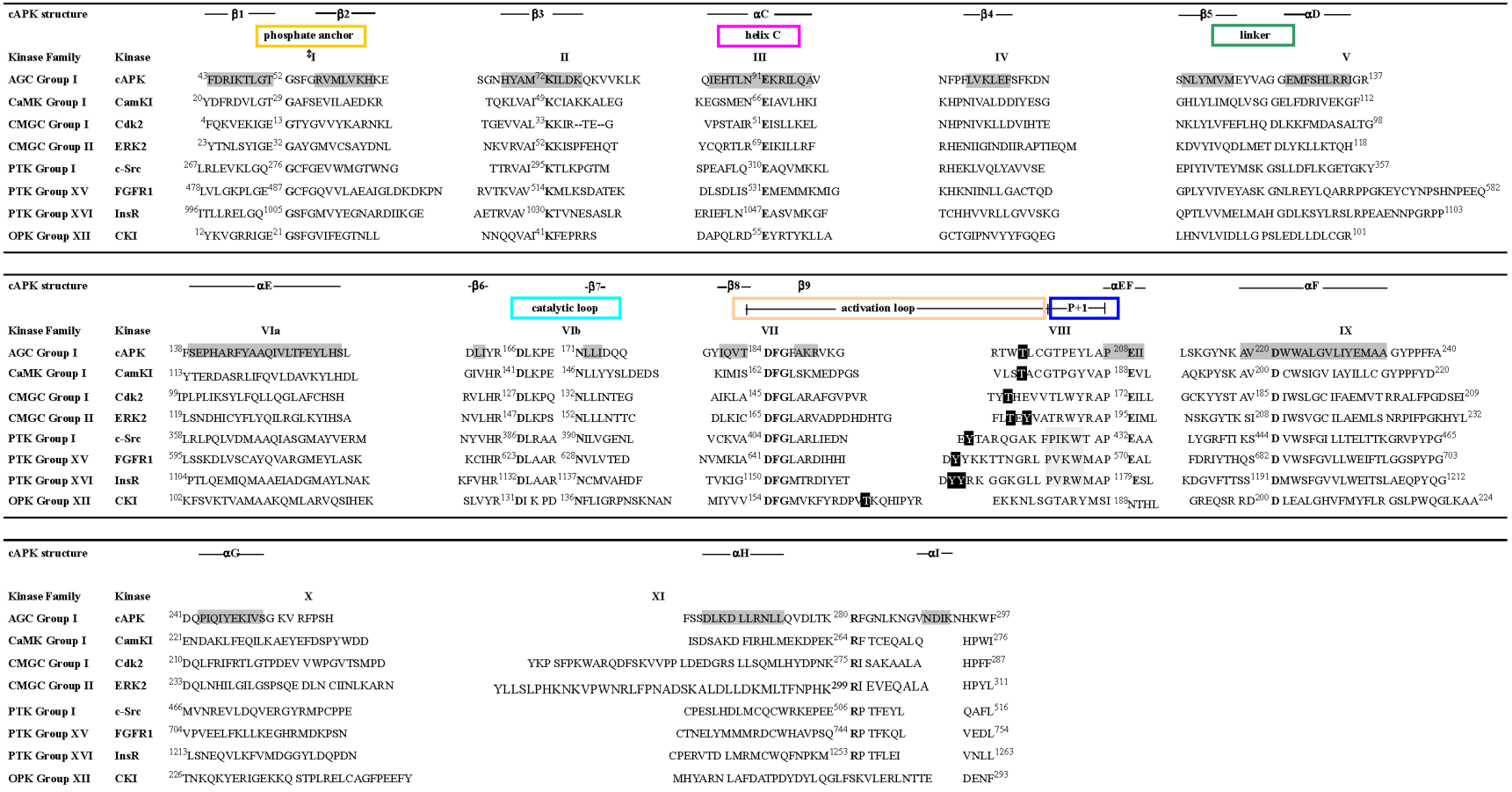
3.3.2 Structural analysis of cAMP-dependent protein kinase (PKA) and major classes of protein kinases [Sowadski and Epstein, 2000]

Three-dimensional structures of the catalytic domains of serine/threonine and protein tyrosine kinase family members have been solved by protein crystallography. Comparison of both structural and sequence data of these kinases allows us to define some common structural elements for all kinases. Figure 3.3 illustrates a representative from each subfamily for which structural data is available.

Protein kinase A (PKA) is the first kinase for which complete backbone NMR assignment was achieved in our laboratory [Langer et al., 2004]. PKA is the first kinase for which crystal structure (catalytic domain) was solved [Knighton et al., 1991], since then all the kinase structures solved till date carry the same architecture, consisting of two lobes linked by a short peptide. The upper lobe consists of five antiparallel β-strands and helix C. The lower lobe consists of the substrate binding site and is predominantly α-helical. ATP binds in the deep cleft between the lobes. Figure 3.4 shows ribbon diagram of cAPK and its various subdomains. The overall topology of the protein kinase catalytic domain, extending from β-strand 1 through helix H, is highly conserved (except helices A and B) among all other reported structures of protein kinases and can be divided into the following subdomains:

- **Phosphate anchor:** Helix A is connected to β-strand 1, followed by the phosphate anchor, encompassing the signature motif Gly50-X-Gly52-X-XGly55. The phosphate anchor forms several hydrogen bonds with oxygens of the β- and γ-phosphates of ATP and is followed by β-strand 2.

- **Nucleotide binding motif:** β-strand 2 is followed by β- strand 3, carrying invariant Lys72, which binds oxygens of the α- and β-phosphates of ATP. Three antiparallel β-strands create the unique fold of the nucleotide binding site of



cAPK structure		β1 — β2 (phosphate anchor)	β3	αC (helix C)	β4	β5 — αD (linker)
Kinase Family	Kinase	I	II	III	IV	V
AGC Group I	cAPK	43FDRIKTLGT52 GSFGRVMLVKHKE	SGNHYAM72KILDKQKVVKLK	QIEHTLN91EKRILQAV	NFPFLVKLEFSFKDN	SNLYMVMEYVAG GEMFSHLRRIGR137
CaMK Group I	CamKI	20YDFRDVLGT29 GAFSEVILAEDKR	TQKLVAI49KCIAKKALEG	KEGSMEN66EIAVLHKI	KHPNIVALDDIYESG	GHLYLIMQLVSG GELFDRIVEKGF112
CMGC Group I	Cdk2	4FQKVEKIGE13 GTYGVVYKARNKL	TGEVVAL33KKIR--TE--G	VPSTAIR51EISLLKEL	NHPNIVKLLDVIHTE	NKLYLVFEFLHQ DLKKFMDASALTG98
CMGC Group II	ERK2	23YTNLSYIGE32 GAYGMVCSAYDNL	NKVRVAI52KKISPFEHQT	YCQRTLR69EIKILLRF	RHENIIGINDIIRAPTIEQM	KDVYIVQDLMET DLYKLLKTQH118
PTK Group I	c-Src	267LRLEVKLGQ276 GCFGEVWMGTWNG	TTRVAI295KTLKPGTM	SPEAFLQ310EAQVMKKL	RHEKLVQLYAVVSE	EPIYIVTEYMSK GSLLDFLKGETGKY357
PTK Group XV	FGFR1	478LVLGKPLGE487 GCFGQVVLAEAIGLDKDKPN	RVTKVAV514KMLKSDATEK	DLSDLIS531EMEMMKMIG	KHKNIINLLGACTQD	GPLYVIVEYASK GNLREYLQARRPPGKEYCYNPSHNPEEQ582
PTK Group XVI	InsR	996ITLLRELGQ1005 GSFGMVYEGNARDIIKGE	AETRVAV1030KTVNESASLR	ERIEFLN1047EASVMKGF	TCHHVVRLLGVVSKG	QPTLVVMELMAH GDLKSYLRSLRPEAENNPGRPP1103
OPK Group XII	CKI	12YKVGRRIGE21 GSFGVIFEGTNLL	NNQQVAI41KFEPRRS	DAPQLRD55EYRTYKLLA	GCTGIPNVYYFGQEG	LHNVLVIDLLG PSLEDLLDLCGR101

cAPK structure		αE	-β6- -β7- (catalytic loop)	β8 — β9 (activation loop)	αEF — P+1	αF
Kinase Family	Kinase	VIa	VIb	VII	VIII	IX
AGC Group I	cAPK	138FSEPHARFYAAQIVLTFEYLHSL	DLIYR166DLKPE 171NLLIDQQ	GYIQVT184DFGFAKRVKG	RTWTLCGTPEYLAP208EII	LSKGYNKAV220DWWALGVLIYEMAA GYPPFFA240
CaMK Group I	CamKI	113YTERDASRLIFQVLDAVKYLHDL	GIVHR141DLKPE 146NLLYYSLDEDS	KIMIS162DFGLSKMEDPGS	VLSTACGTPGYVAP189EVL	AQKPYSKAV200D CWSIGV IAYILLC GYPPFYD220
CMGC Group I	Cdk2	99IPLPLIKSYLFQLLQGLAFCHSH	RVLHR127DLKPQ 132NLLINTEG	AIKLA145DFGLARAFGVPVR	TYTHEVVTLWYRAP172EILL	GCKYYSTAV185D IWSLGC IFAEMVT RRALFPGDSEI209
CMGC Group II	ERK2	119LSNDHICYFLYQILRGLKYIHSA	NVLHR147DLKPS 152NLLLNTTC	DLKIC165DFGLARVADPDHDHTG	FLTEYVATRWYRAP195EIML	NSKGYTKSI208D IWSVGC ILAEMLS NRPIFPGKHYL232
PTK Group I	c-Src	358LRLPQLVDMAAQIASGMAYVERM	NYVHR386DLRAA 390NILVGENL	VCKVA404DFGLARLIEDN	EYTARQGAK FPIKWT AP432EAA	LYGRFTI KS444D VWSFGI LLTELTTKGRVPYPG465
PTK Group XV	FGFR1	595LSSKDLVSCAYQVARGMEYLASK	KCIHR623DLAAR 628NVLVTED	NVMKIA641DFGLARDIHHI	DYYKKTTNGRL PVKWMAP570EAL	FDRIYTHQS682D VWSFGVLLWEIFTLGGSPYPG703
PTK Group XVI	InsR	1104PTLQEMIQMAAEIADGMAYLNAK	KFVHR1132DLAAR1137NCMVAHDF	TVKIG1150DFGMTRDIYET	DYYRK GGKGLL PVRWMAP1179ESL	KDGVFTTSS1191DMWSFGVVLWEITSLAEQPYQG1212
OPK Group XII	CKI	102KFSVKTVAMAAKQMLARVQSIHEK	SLVYR131DI K PD 136NFLIGRPNSKNAN	MIYVV154DFGMVKFYRDPVTKQHIPYR	EKKNLSGTARYMSI188NTHL	GREQSR RD200D LEALGHVFMYFLR GSLPWQGLKAA224

cAPK structure		αG	αH	αI
Kinase Family	Kinase	X	XI	
AGC Group I	cAPK	241DQPIQIYEKIVSG KV RFPSH	FSSDLKD LLRNLLQVDLTK280 RFGNLKNGVNDIKNHKWF297	
CaMK Group I	CamKI	221ENDAKLFEQILKAEYEFDSPYWDD	ISDSAKD FIRHLMEKDPEK264 RF TCEQALQ	HPWI276
CMGC Group I	Cdk2	210DQLFRIFRTLGTPDEV VWPGVTSMPD	YKP SFPKWARQDFSKVVPP LDEDGRS LLSQMLHYDPNK275 RI SAKAALA	HPFF287
CMGC Group II	ERK2	233DQLNHILGILGSPSQE DLN CIINLKARN	YLLSLPHKNKVPWNRLFPNADSKALDLLDKMLTFNPHK299 RI EVEQALA	HPYL311
PTK Group I	c-Src	466MVNREVLDQVERGYRMPCPPE	CPESLHDLMCQCWRKEPEE506 RP TFEYL	QAFL516
PTK Group XV	FGFR1	704VPVEELFKLLKEGHRMDKPSN	CTNELYMMMRDCWHAVPSQ744 RP TFKQL	VEDL754
PTK Group XVI	InsR	1213LSNEQVLKFVMDGGYLDQPDN	CPERVTD LMRMCWQFNPKM1253 RP TFLEI	VNLL1263
OPK Group XII	CKI	226TNKQKYERIGEKKQ STPLRELCAGFPEEFY	MHYARN LAFDATPDYDYLQGLFSKVLERLNTTE	DENF293

Kinase family and domain classification are based on the nomenclature of Hanks and Hunter (Hanks and Hunter, 1995). Secondary structure and functional domain nomenclature is adopted from the first protein kinase crystal structure, cAMP dependent protein kinase catalytic subunit (cAPK) (Knighton *et al.*, 1991a, b). Descriptions of these conserved structural regions are presented more fully in the text. β Strand and helical regions are shaded grey. The upper panel refers to the upper domain. The two lower panels refer to the lower domain. Invariant residues are in bold. Known, key phosphorylation sites are outlined in black. ERK2 protein kinase contains two sites, Thr183 and Tyr185, both subject to phosphorylation by the dual-specificity protein kinase MEK1. Sequence PxxWx in domain VIII (shaded box) is conserved in every protein tyrosine kinase. It is responsible for dictating the specificity of protein tyrosine kinases versus Ser/Thr protein kinases.

Figure 3.3: Protein kinase sequence and structural elements. Sequence alignment of catalytic cores of protein kinases for which crystal structures have been determined [Sowadski and Epstein, 2000].

protein kinases. β-Strand 3 is followed by helix B (present only in cAPK), helix C, and β-strands 4 and 5.

- **Helix C:** In most of the protein kinase structures helix C shows the largest displacement. The salt bridge between the Glu91 and Lys72 is also conserved in most kinases, except for the inactive cyclin dependent protein kinase (Cdk2) structure, but present in the crystal structure of its complex with an activator cyclin A.

- **Linker:** The N- and C- lobe are connected by a linker. This portion of the structure provides two main-chain hydrogen bonds to the 6-amino and N1 positions of the ATP adenine ring. The last residue of the linker Glu127 forms a hydrogen bond with the 2'OH of ATP. Interesting fact is that the ϕ and θ of the main chain of this linker region vary among different protein kinases, reflecting displacement in the upper domain with respect to the lower. This region also gives the conformational plasticity for the nucleotide binding event.

- **Catalytic loop:** This loop connects the β-strands 6 and 7 and harbors the crucial set of residues required for the catalytic activity of the enzyme, starting with Tyr164 and Arg165. Tyr164 forms a hydrogen bond with the conserved Asp220 and Arg165 present in most of protein kinases, provides two hydrogen bonds to the oxygen of the Thr197 phosphate. The conserved catalytic base Asp166 and Asn171, the ligand to one of the metal sites, are located within this loop. Ser/Thr and Tyr kinases show sequence diversity in this region. In Ser/Thr kinases the Lys168 interacts with the γ-phosphate of ATP during catalysis. In Tyr kinases this role of Lys is replaced by Arg, exemplified by Arg1136 in the crystal structure of Insulin receptor tyrosine kinase (InsR).

- **Activation loop:** The catalytic loop and β-strand 7 are followed by β-strand 8 and the conserved Asp184-Phe185-Gly186 (DFG) motif. The conserved Asp184 forms co-ordinate complex with metal ion (Mg^{2+}). This motif is followed by β-strand 9 and the activation loop with Thr197. Phosphorylation of Thr or Tyr in the activation loop is often critical for activation of many protein kinases. This is central for binding substrates with Ser or Tyr containing protein kinases. The phosphorylation state of the activation loop is critical for the substrate binding. Followed by the activation loop is the P+1 loop (P= substrate P (Phosphorylation) site), which has the P+1 site of the substrate, a critical specificity determinant for substrate recognition. The sequence PxxWx

in this domain is conserved in tyrosine kinases and is responsible for specificity of PTK versus Ser/Thr protein kinases.

- **Structural pair Glu208-Arg280:** The stabilization and positioning of the P+1 loop with respect to the C-lobe of protein kinase is brought by Glu208-Arg280 pair.
- **C-terminal domain:** This part of the kinase is composed of the helix F-J. This part of the molecule undergoes large motional changes during the binding events.

All of the above structural features are summarized pictorically in Figure 3.4.

3.3.3 Substrate specificity: How does a kinase recognize its substrate?

All kinases transfer γ-phosphate from ATP either to Ser/Thr/Tyr of the substrate, and hence recognition of specific substrate is vital for efficient signal transduction. This recognition is accomplished at several stages viz.,

1. colocalization of the kinase and substrate in the same cell compartment
2. protein-protein interaction module
3. primary amino acid sequences near to and distant from the phosphorylation site (P site).

3.3.4 Mechanism of regulation of protein kinase activity

Protein kinases play a vital and important role in the cellular signal transduction process. The switching ON/OFF of various protein kinases is vital for the proper functioning of the cell, this is regulated by various mechanisms viz.,

- **Regulation by subunits:** cAPK in its inactive state exists as a tetrameric holoenzyme complex (R_2C_2) with its regulatory subunit (R) and catalytic subunit (C). Binding of cAMP releases the monomeric active cAPK. Other examples include CDK2 by cyclin A.
- **Regulation by domains:** Ca^{2+}/calmodulin-dependent kinase is activated by Ca^{2+}-calmodulin binding to the C-terminal tail of the kinase. In the absence of Ca^{2+}-calmodulin, the tail blocks the substrate and ATP binding site and illustrates a autoinhibitory mechanism.

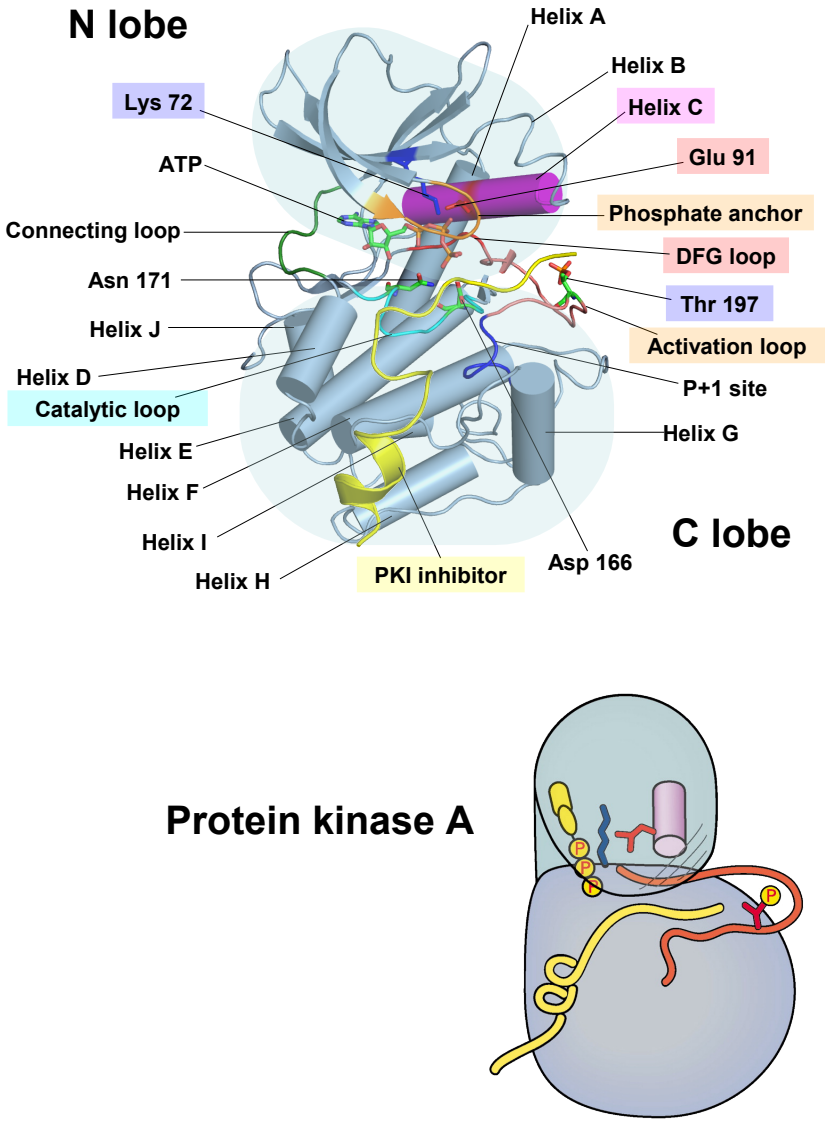


Figure 3.4: Structural architecture of a typical protein kinase. The Catalytically Active Conformation of the Protein Kinase A(Top) The crystal structures of protein kinase A (PDB:1ATP, PKA, a serine/threonine kinase) [Zheng et al., 1993]. Key structural elements within the kinase domain are colored and labeled. The PKI peptide inhibitor for PKA is colored yellow. Bound nucleotide and several absolutely conserved residues within the active site are indicated. (Bottom) The above structure has been schematized to clearly depict the conformational transitions involved in the regulation of kinase activity, with particular emphasis on the helix C and activation loop. The catalytic lysine and glutamate (Lys72 and Glu91 in PKA) are shown schematically. The schematic representation reproduced from Huse and Kuriyan [2002].

- **Regulation by phosphorylation of the activation loop:** Phosphorylation of the Ser/Thr/Tyr residues within the activation loop of many kinases results in conformational changes in the protein leading to activated state of the protein. These include Cdk2, ERK2, src, hck, lck etc.

3.3.5 Oncogenic kinases in cancer

A cell can follow many paths to malignancy. The origin of cancer lies in the genetic material of the cell. Up to 20 % of the 32,000 human genes sequenced and among these are more than 520 kinases [Blume-Jensen and Hunter, 2001]. Most of cellular signalling involves these kinases and their deregulation in human cancer can initiate or alter signals that eventually lead to cell proliferation and transformation. Transmembrane (e.g. EGFR, PDGFR) or cytoplasmic (Src, Abl) tyrosine kinases are found mutated in a variety of human tumors. Cytoplasmic serine threonine kinases (Raf, Akt, Tpl-2) are also mutated or activated in several types of human malignancies. Kinases transduce signals that lead to cell proliferation or inhibition of programmed cell death by activating transcription factors (e.g. AP1, NFkappaB, Myc), inhibiting pro-apoptotic molecules (e.g. Bad, Bax), or they participate in deregulating the cell cycle control. The involvement of kinases in vital signaling events renders them valuable targets for therapeutic intervention in human cancers [Blume-Jensen and Hunter, 2001].

3.3.6 NMR studies on protein kinases

"A picture is worth a thousand words." While this is true undoubtedly for all the protein structures that have been solved by either X-ray crystallography or NMR spectroscopy, but these static three-dimensional structures does not clearly explain the functional biological enzymatic reactions nor is useful for rational drug design [Kay, 1998]. Specially in case of protein kinases which has a large similarity in amino acid sequence, the three-dimensional structures provide important snapshots of different states, but fails to provide the underlying much important dynamic process. The specificity of their function is mostly dictated by small differences in the primary sequence and subtle changes in the overall conformation of the protein occurring as a result of activation by phosphorylation or substrate binding. Studying this dynamic process at atomic resolution with respect to time is the key for elucidating a clear mechanism for biological function, protein engineering and rational drug design.

NMR allows monitoring of a wide variety of such motional process [Jarymowycz and Stone, 2006]. Methods like relaxation analysis, residual dipolar couplings has been used to study these challenging protein kinases, as the catalytic domains become accessible for NMR.

Table 3.2 provides a comprehensive list of protein kinases (either free or in complex) studied by various parameters of NMR spectroscopy viz., backbone assignment (a prerequisite for studying dynamic process by NMR), relaxation studies, and residual dipolar couplings.

Table 3.2: Comprehensive list of published NMR backbone dynamic studies on various domains of human protein kinases

NMR experiment	Protein	Form (free/bound)	Reference
Complete or partial backbone assignment of human protein kinases	ABL kinase	Bound to inhibitors (imatinib, nilotinib, and dasatinib)	[Vajpai et al., 2008]
	cAMP-binding domain A of the PKA regulatory subunit	free	[Esposito et al., 2006]
	mTOR domain responsible for rapamycin binding	free	[Veverka et al., 2006]
	F-actin binding domain of human Bcr-Abl/c-Abl	free	[Wiesner et al., 2005]
	Apo and peptide bound state of SH2 domain of Rous sarcoma viral protein Src	Free and bound (peptide)	[Taylor et al., 2005]
	Mitogen-activated protein (MAP) kinase p38	free	[Vogtherr et al., 2005]
	SH2 domain of human feline sarcoma oncogene FES	free	[Scott et al., 2004]
	Catalytic subunit of cAMP-dependent protein kinase A	free	[Langer et al., 2004]
	SH3-SH2 domain pair from the human tyrosine kinase	free	[Schweimer et al., 2003]
	Partial assignment of cAbl kinase	free	[Strauss et al., 2003]
	Pleckstrin homology domain of human protein kinase B (PKB/Akt)	free	[Auguin et al., 2003]
	SH2 domain of the Csk (C-terminal Src kinase) homologous kinase	free	[Mulhern et al., 2002]
	SH3 domain of phosphatidylinositol 3-kinase (PI3K) and D21N-mutant	free	[Okishio et al., 2001]
	Hck SH2 domain	free	[Zhang et al., 1997]
	SH3 domain of human p56 Lck tyrosine kinase	free	[Hiroaki et al., 1996]
	SH2 domain from Fyn tyrosine kinase	free	[Pintar et al., 1996]
	SH2-SH3 domain of Abl kinase	free	[Gosser et al., 1995]
Relaxation studies of fast backbone dynamics in proteins	Receptor tyrosine kinase ephrin-BS ectodomain	free	[Ran et al., 2008]

Table 3.2: (continued...)

NMR experiment	Protein	Form (free/bound)	Reference
	ABL kinase	Bound to inhibitors (imatinib, nilotinib, and dasatinib)	[Vajpai et al., 2008]
	Receptor tyrosine kinase ErbB2	free	[Bocharov et al., 2008]
	MARK3-UBA domain	free	[Murphy et al., 2007]
	Phosphatidylinositol 3-kinase (PI3K) SH3 domain	free	[Ahn et al., 2006]
	SH3 domains of drkN and mutant Fyn	free	[Bezsonova et al., 2006]
	Pleckstrin binding domain of the human protein kinase B (PKB/AKT)	free	[Auguin et al., 2004]
	Abl SH3 domain and crk SH2 domain	Free	[Donaldson et al., 2002]
	Backbone motions in ligand binding to c-Src SH3 domain	Free and bound	[Wang et al., 2001]
Residual dipolar couplings (RDCs)	SH2 domain of Syk protein-tyrosine kinase	Free	[Zhang et al., 2008b]
	ABL kinase	Bound to inhibitors (imatinib, nilotinib, and dasatinib)	[Vajpai et al., 2008]
	p38α MAP kinase-SB203580	Free and bound	[Honndorf et al., 2008]
	Coiled-coil domain of cGMP dependent protein kinase Iα	free	[Schnell et al., 2005]
	SH3-SH2 domain orientation in Src kinase Fyn	free	[Ulmer et al., 2002]

3.4 Heat shock protein of 90 kDa (Hsp90)-a kinome chaperone

Cancer cells are characterized by unregulated cell division often promoted by protein kinases. Therefore, protein kinases are often targeted by small molecule inhibitors, albeit with the problem of insufficient target selectivity due to the high sequence similarity between different kinase family members. As alternative route, drugs that inhibit the molecular chaperone Hsp90 are currently in clinical trials due to their ability to promote degradation of many kinases. In this section we will review more about Heat Shock Proteins in general and more about Hsp90, and also address how its function in cancer cells differs from healthy cells.

3.4.1 Introduction

Cells under normal physiological conditions are under perfect state of equilibrium, which allows it to function normally and is termed as homeostasis. Exposure of these cells in most tissues to environmental stress inducers (viz., heat, heavy

metals, hypoxia and acidosis) causes an alteration in the homeostatic condition of the cells. In order to maintain homeostasis, cells produce a small group of proteins that are collectively called as 'heat-shock' or stress proteins [Leppä and Sistonen, 1997]. Most of these heat shock proteins are actually proteins that prevent improper associations and assist in the correct folding and maturation of other cellular proteins and there by called as 'molecular chaperones' that guard the cell. Their basal level expression helps the normal protein folding and guards the cell from the dangers of protein misfolding and aggregation [Wegele et al., 2004]. Their increased expression is an adaptive response to enhance cell survival. Both these functions are needed in tumors and reflect the fact how malignant cells are able to survive in hostile environment. These chaperones also allow the tumor cells to tolerate mutations in the key signaling proteins, which is lethal and drive oncogenesis [Jolly and Morimoto, 2000]. Chaperones also impart many other physiological functions [Whitesell and Lindquist, 2005] and are schematically summarized in Figure 3.5.

The molecular chaperone Heat shock protein of 90 kDa (Hsp90) in particular has been identified as a potential anti-cancer target as it allows mutant proteins to retain function hence allowing the cancer cells to tolerate the lacunae in signaling created by such oncoproteins [Whitesell and Lindquist, 2005]. Hsp90 is an abundant and essential protein, accounting for 1-2 % of the total protein amount in unstressed eukaryotic cells. A 2 to 10 fold higher expression level of Hsp90 is found in cancer cells and virally transformed cells [Ferrarini et al., 1992; Welch and Feramisco, 1982; Yufu et al., 1992], suggesting a crucial role of Hsp90 for growth and/or survival of tumor cells. Hsp90 is a protein chaperone that helps a variety of client proteins to adopt their native conformation. Several co-chaperones such as Hop/Sti1, FKBP52, cyclophilin 40, and p50/Cdc37 are identified to be associated with Hsp90. These co-chaperones are required for binding to diverse client proteins and controlling the ATPase cycle of Hsp90 [Picard, 2002; Terasawa et al., 2005; Wegele et al., 2004]. The important components of Hsp90-based chaperone machinery are summarized in Table 3.3. The ATP-binding site of Hsp90 is located in the N-terminal domain. Low molecular weight compounds that interfere with ATP-binding such as geldanamycin or radicicol inhibit the Hsp90 function by competing for interaction site.

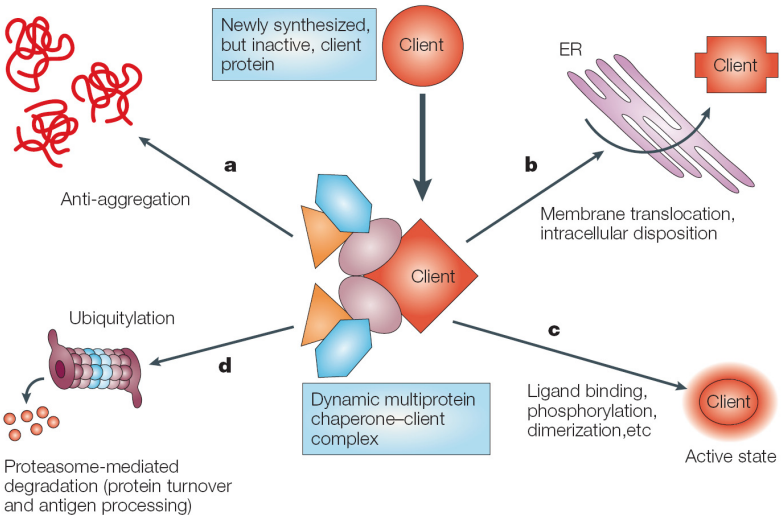


Figure 3.5: Role of molecular chaperones in regulating protein homeostasis. Newly synthesized, conformationally labile client proteins associate with multi-protein complexes that contain various chaperones, co-chaperones and accessory molecules (different colored shapes). The particular components of a complex vary according to the client and also help specify the function of a particular complex. Dynamic association of a client with chaperone complexes can prevent its aggregation (a) and assist in its intracellular trafficking, especially its translocation across membranous structures such as the endoplasmic reticulum (ER) (b). For many clients involved in signal transduction pathways, association with the chaperone machinery maintains the protein in a meta-stable state that allows it to be activated by specific stimuli such as ligand binding, phosphorylation or assembly into multisubunit signalling complexes (c). In the absence of appropriate stimuli, chaperone complexes can target the client for degradation through the ubiquitin–proteasome pathway, thereby regulating its steady-state cellular level (d). Figure reproduced from Whitesell and Lindquist [2005].

3.4.2 Chaperone alteration in cancer

Both solid tumors and haematological malignancies are characterized by increased level of expression of HSPs [Ciocca et al., 1993; Conroy et al., 1998; Kaur and Ralhan, 1995; Kimura et al., 1993; Ralhan and Kaur, 1995; Santarosa et al., 1997]. Hsp90 and its co-chaperones also modulate tumor cell apoptosis. Many of the client proteins of Hsp90 are crucial elements of the cellular signal transduction pathway. A sample of the many Hsp90 client proteins that are known to be involved in cancer, and organized according to their contributions to the malignant phenotype, is provided in Table 3.4.

Protein kinases represent the largest class of Hsp90 client proteins and many have been shown to cause cancer when deregulated. These protein kinases depend on Hsp90 for their proper folding and functioning. The first protein kinase shown

Table 3.3: Important components of Hsp90 chaperone machinery

Protein Family	Classification	Function
Hsp90	Chaperone	Supports meta-stable protein conformations, especially in signal transducers
Hsp70	Chaperone	Helps fold nascent polypeptide chains; participates in assembly of multiprotein complexes
Hsp40	Co-chaperone	Stimulates Hsp70 ATPase activity
HIP (Hsp70-interacting protein), HOP (Hsp70/Hsp90-organizing protein)	Adapters	Mediate interaction of Hsp90 and Hsp70
Cdc37 (cell division cycle protein 37)/p50	Co-chaperone	Modulates interaction with kinases
Aha1 (activator of Hsp90 ATPase activity)	Co-chaperone	Stimulates Hsp90 ATPase activity
p23	Co-chaperone	Stabilizes Hsp90 association with clients
Immunophilin	Prolylisomerase	Modulates interactions with hormone receptors

Table 3.4: Hsp90 clients and the malignant phenotype

Phenotype	Clients
Uncontrolled proliferation	Receptor tyrosine kinases, serine/threonine kinases, steroid hormone receptors
Immortalization	Telomerase
Impaired apoptosis	AKT
Angiogenesis	HIFIα (hypoxia-inducible factor 1α)
Invasion/metastasis	MMP2 (matrix mettaloprotease 2)

to interact with Hsp90 was v-Src tyrosine kinase. Since then, numerous other kinases [viz., Tyrosine kinases (v-Src, c-Src, Lck, c-Fgr, Insulin receptor, etc), Serine-threonine kinases (v-Raf, c-Raf, Gag-Mil, MEK, CDK4, etc) have been reported to form stable complexes with Hsp90. In tumor cells, the dynamic low affinity interactions of Hsp90 with its client proteins such as kinases maintain them in a latent but readily activated state. The conformationally unstable oncogenic mutants of such clients require more of Hsp90 to stabilize them, hence justifying the over expression of Hsp90 in tumors, suggesting that Hsp90 itself is a valid anti-cancer target of pharmaceutical interest [Whitesell and Lindquist, 2005].

3.4.3 Hsp90 structure and function

Hsp90 primarily resides in the cytoplasm of the cell. In vertebrates two isoforms, Hsp90α and Hsp90β, which are 76% identical and are the consequences of a gene duplication about 500 million years ago [Krone and Sass, 1994; Moore et al., 1989]. Hsp90β is constitutively expressed and was also frequently denoted as Hsp86 and Hsp84. In contrast, Hsp90α is inducible and is slightly smaller than Hsp90β. The functional differences between these two isoforms is poorly understood [Sreedhar et al., 2004]. Apart from this, recently a membrane-associated variant Hsp90N has been reported with limited details of its cellular functions [Grammatikakis et al., 2002].

Hsp90 predominantly exists as a homodimer in the cytoplasm. Each of these homodimer is made of three domains, which have important functional interactions, as summarized in Figure 3.6.

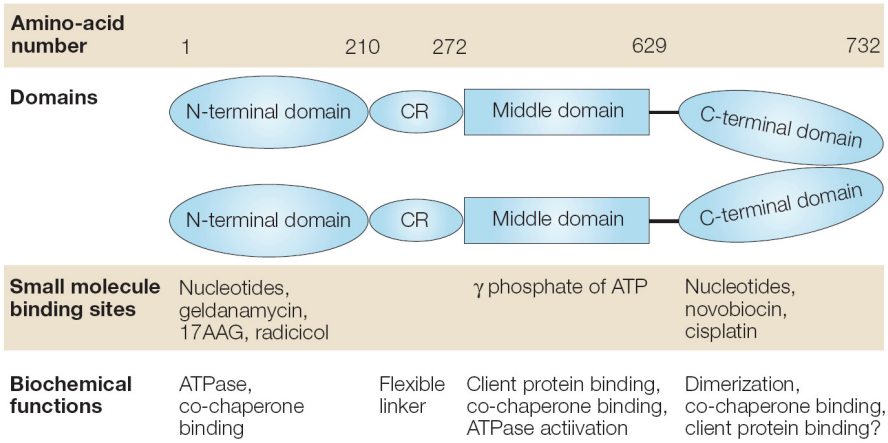


Figure 3.6: Schematic representation of the Hsp90 dimer. The numbering 1-732 indicates the approximate positions in the amino acid sequence of the human protein that define its functional domains. 'CR' refers to a charged region, which serves as a flexible linker between the N-terminal and middle domains. The locations where various small molecules bind HSP90 (heat-shock protein of 90 kDa) and modulate its function are indicated (17AAG, 17-allylaminogeldanamycin; GA, geldanamycin). The biochemical functions of each domain are also indicated. Figure reproduced from Whitesell and Lindquist [2005].

The crystal structure of full length yeast Hsp90 was recently reported(Ali et al., 2006). The N-terminal (both human and yeast) [Prodromou et al., 1997; Stebbins et al., 1997] and middle domain (yeast) structures has also been reported [Meyer et al., 2003].

The N-terminal domain of approximately 25 kDa, contains an unusual adenine-nucleotide-binding pocket known as the Bergerat fold [Dutta and Inouye, 2000]. This domain is mainly responsible for the hydrolysis of ATP to ADP and plays an essential role in the chaperoning activity of the Hsp90 dimer, for which two models (yeast Hsp90 and human Hsp90) has been proposed [McLaughlin et al., 2004; Meyer et al., 2003]. The model proposed after involving a study with yeast Hsp90 indicates dimerization of N-terminal domain of Hsp90 upon ATP binding, whereas there was no evidence found for dimerization in human Hsp90. Cdc37 a co-chaperone of Hsp90, binds to the N-terminal domain and plays an important role in recruiting kinases to the Hsp90 [Roe et al., 2004]. Disrupting the association of Cdc37 to Hsp90 by designing small molecule inhibitors has been recently proposed for increasing the selectivity in targeting oncogenic kinases [Gray et al., 2008; Smith et al., 2009; Smith and Workman, 2009].

A charged linker connects the N-terminal domain of Hsp90 to its middle domain, of approximately 35 kDa. Structure of this middle domain indicates an important role of harboring the γ-phosphate of ATP molecules that bind to the N-terminal pocket of Hsp90 [Meyer et al., 2003; Söti et al., 2002]. Additionally, AHA1, a co-chaperone of Hsp90, promotes association of the N- and the middle domain and accelerates the rate of ATP-hydrolysis [Meyer et al., 2004; Panaretou et al., 2002].

Finally, another flexible linker connects the middle domain to the C-terminal domain, of approximately 12 kDa. This domain is mainly responsible for the dimerization, and whose deletion drastically impairs the ATPase activity of Hsp90, providing a perfect example for intermolecular and intramolecular interactions are required for enzymatic reaction. This domain also carries a conserved EEVD motif responsible for the binding of various client proteins containing tetratricopeptide containing repeats (TPR)-domain such as immunophilins, HOP and protein phosphatase 5 (PP5), which increase the specificity of Hsp90 complexes [Prodromou et al., 1999; Scheufler et al., 2000].

3.4.4 Inhibition of Hsp90 function

Hsp90 plays a key role in the conformational maturation of oncogenic signalling proteins, including HER-2/ErbB2, Akt, Raf-1, Bcr-Abl and mutated p53 [Kamal et al., 2003]. ATP-binding pocket residing in the N-terminal domain of Hsp90 is the binding site of several inhibitors of Hsp90, viz., geldanamycin, 17AAG and

radicicol. Several other Hsp90 binding drugs are summarized in Table 3.5. The drug-bound Hsp90 sensitises its client proteins to recruit E3 ubiquitin ligases such as CHIP (carboxy-terminus of Hsp70-interacting protein), which otherwise would have been stabilized by Hsp90-containing chaperone complexes. This recruitment leads to increased proteasome-mediated degradation of the clients and depletion of their cellular levels.

Table 3.5: Hsp90-binding drugs

Binding site	Chemical class	Selected examples
N-terminal ATP-binding pocket	Benzoquinone ansamycin	Geldanamycin (GA), 17AAG (17-(demethoxy).17-allylamino geldanamycin), 17DMAG (17-(demethoxy).17-dimethyl-aminoethylamino geldanamycin)
N-terminal ATP-binding pocket	Macrolide	Radicicol and related oxime derivatives, β-zearalenol
N-terminal ATP-binding pocket	Purine scaffold	PU24FC1
N-terminal ATP-binding pocket	Pyrazole	CCT018159
N-terminal ATP-binding pocket	Hybrid	Radamycin, GA dimmer, GA-testosterone, GA-estrogen
C-terminus	Novioylcoumarin crosslinker	Novobiocin, coumermycin, cis-platin
Unknown	Histone deacetylase inhibitor	Depsipeptide

Clinical and preclinical observations have shown that Hsp90 inhibitors at therapeutic doses significantly impair tumor growth with minimal effects on the normal tissues. However it is now becoming more clear that the currently available drugs also alter some of the normal functions of the cells [Whitesell and Lindquist, 2005]. Selectivity of the drugs towards the tumor cells in comparison to normal tissues come from the fact that Hsp90 has an enhanced ATPase activity in tumor cells, since it mainly exists as a multi-chaperone complex, when compared to its mostly latent state in normal cells [Kamal et al., 2003].

3.5 Cell division cycle protein 37 (Cdc37)-a kinome co-chaperone

In the previous section, we have seen that one of the major functions of Hsp90 is stabilization of protein kinases. Hsp90 does not directly bind to protein kinases, but this interaction is mediated by the co-chaperone Cdc37/p50. Hence, Cdc37 has also been dubbed as a kinase targeting subunit of the Hsp90 chaperone machinery

[Hunter and Poon, 1997]. Interestingly, Cdc37 is also capable to act as a chaperone by itself independent of Hsp90 [MacLean and Picard, 2003]. Hsp90-Cdc37 complex is considered to mediate carcinogenesis by stabilizing a variety of different oncogenic kinases in malignant cells. In this section, we will review more about Cdc37 in general and more about Cdc37-Hsp90 interaction, and also address potential implications of Cdc37 as drug target.

3.5.1 Introduction

The cell division cycle protein 37 (Cdc37) was first identified in *S. cerevisiae* as a protein required for the 'Start' event in the cell division cycle [Reed, 1980]. In yeast, *cdc37* mutations cause synthetic lethality (as defined by [Kaelin, 2005]), in combination with thermosensitive *cdc28* or *kin28* mutations [Reed et al., 1985; Valay et al., 1995]. The mammalian homolog of Cdc37, also referred to as $p50^{cdc37}$, was first identified within a complex consisting of the Rous sarcoma virus encoded oncogene $pp60^{v\text{-}scr}$ and Hsp90 [Brugge, 1986; Whitelaw et al., 1991]. Further studies confirmed that protein kinases are the most favored targets of Cdc37 [Mort-Bontemps-Soret et al., 2002]. A plethora of other oncogenic kinases including Raf isoforms, Akt, and ErbB2 have been identified as potential clients of Cdc37 (Figure 3.7) [Basso et al., 2002; Grammatikakis et al., 1999; Lamphere et al., 1997]. Expression of Cdc37 is upregulated in cancer cells and tissues and can promote tumorigenesis when overexpressed [Schwarze et al., 2003; Stepanova et al., 2000a].

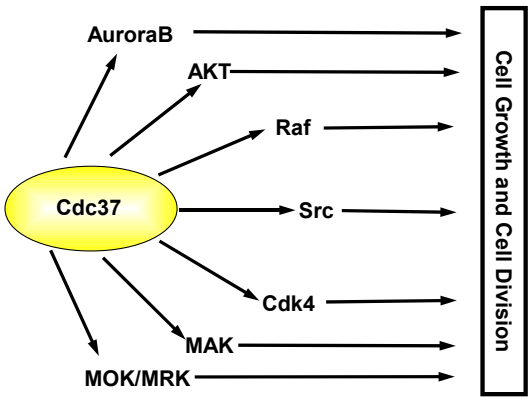


Figure 3.7: Cdc37 is essential for folding and stabilization of many kinases. This list of kinases also include several of the oncogenic kinases like AKT, Raf, Aurora, etc., which otherwise are also required for normal cellular function.

Table 3.6 summarizes most of the client proteins of Cdc37 in association with Hsp90. Further in yeast it has been shown that Cdc37 can function alone in kinase biogenesis [Mandal et al., 2007]. Apart from this Cdc37 also interacts with several non-kinase proteins like androgen receptor (for proper signaling), viral receptor

transcriptase, the transcription factor MyoD and nitric oxide synthase [Harris et al., 2006; Yun and Matts, 2005]. All these can be summarized into the fact that Cdc37 essentially has two roles; in one role it acts like a bridge between kinase and Hsp90 and modulates the ATPase activity of Hsp90 during the process of maturation of kinases; and in the second role in which Cdc37 functions in an Hsp90 independent manner.

Table 3.6: Cdc37 interacting proteins (other than Hsp90)*

Clients	Interaction	Evidence	Reference
AKT (Protein Kinase B)	Biochemical	Endogenous Akt IPs with Hsp90 and Cdc37. Cdc37 can bind Akt directly, not affected by Hsp90 inhibitors.	Basso et al. (2002) J. Biol. Chem. 277: 39858
Androgen receptor	Genetic	Yeast *cdc37-34* has decreased activation of heterologous AR	Rao et al.(2001) J. Biol. Chem. 276: 5814
	Biochemical	Human Cdc37 binds to ligand-binding domain of AR in vitro, partially Hsp90dependent. N-terminal domain (1-173) is dominant negative	Fliss et al. (1997) Mol. Biol. Cell 8: 2501
Aurora B	Genetic & Biochemical	Mammalian & Drosophila Cdc37, inactivation of Hsp90 or Cdc37 reduces Aurora B levels, Aurora B or Cdc37 RNAi phenotypes are almost indistinguishable. Co-IP	Lange et al. (2002) EMBO J. 21: 5364
B-Raf	Biochemical	Co-IP	Vaughan et al. (2006) Mol. Cell 23: 697; Grbovic et al. (2006) Proc. Natl. Acad. Sci. USA 103, 57
Cak1	Genetic	Synthetic lethal screen	Mort-Bontemps-Soret et al. (2002) Mol. Genet. Genomics 267: 447
	Genetic	Yeast *cdc37-1* has reduced Cak1 level and activity. Cdc37 stabilises Cak1 after translation	Farrell & Morgan (2000) Mol. Cell. Biol. 20: 749
Cdc2	Genetic & Biochemical	*S. pombe* Cdc2 fails to associate with cyclin in cdc37 mutant strain; genetic interactions and co-IP.	Turnbull et al. (2006) J. Cell Sci. 119: 292
Cdc5	Genetic	Synthetic lethal screen	Mort-Bontemps-Soret et al. (2002) Mol. Genet. Genomics 267: 447
Cdc7	Genetic	Synthetic lethal screen	Mort-Bontemps-Soret et al. (2002) Mol. Genet. Genomics 267: 447
Cdc15	Genetic	Synthetic lethal screen (and kinase assay in S. pombe)	S. cerevisiae: Mort-Bontemps-Soret et al. (2002) Mol. Genet. Genomics 267: 447 / S. pombe (=Cdc7) Liang and Fantes (2007) Euk. Cell 6: 1089
Cdc25c	Biochemical	Co-IP	Garcia-Morales et al. (2007) Oncogene 26: 7185
Cdc28	Genetic	Yeast *cdc37-1* mutant affects Cdc28/cyclin complex formation and function	Gerber et al. (1995) Proc. Natl. Acad. Sci. USA 92: 4651
	Genetic	Yeast Cdc37 stabilises Cdc28 after translation	Farrell & Morgan (2000) Mol. Cell. Biol. 20: 749
	Genetic & Biochemical	Yeast *CDC37* is a multi-copy suppressor of *cdc28*. yeast two-hybrid interaction between Cdc37 C-terminus and Cdc28 N-terminus	Mort-Bontemps-Soret et al. (2002) Mol. Genet. Genomics 267: 447
Cdk2	Biochemical	pull-down assays	Prince et a. (2005) Biochem. 44: 15287
Cdk4	Biochemical	Immunoprecipitation with mammalian Cdc37 and Hsp90	Dai et al. (1996) J. Biol. Chem. 271: 22030
	Biochemical	Mammalian co-IP (preferentially without cyclin) with Hsp90	Stepanova et al. (1996) Genes & Dev. 10: 1491
		CDC37 transgenic mice develop mammary tumours, i.e. can function as an oncogene. Can co-operate with Cyclin D1 or with c-myc in transformation of tissue. See also Schwarze et al. (2003) Cancer Res. 63: 4614.	Stepanova et al. (2000) Mol. Cell. Biol. 20: 4462
	Biochemical	Yeast 2-hybrid with human Cdc37. GST-Cdc37 co-IP	Lamphere et al. (1997) Oncogene 14: 1999
Cdk6	Biochemical	GST-Cdc37 co-IP	Lamphere et al. (1997) Oncogene 14: 1999
Cdk9	Biochemical	co-IP with Cdk9 before assembly with cyclin T1	O'Keeffe et al. (2000) J. Biol. Chem. 275: 279
Cdk11	Biochemical	co-IP	Mikolajczyk and Nelson (2004) Biochem. J. 384: 461
CKII (casein kinase II)	Genetic	Multi-copy suppressor	McCann & Glover (1995) Mol. Biol. Cell (Supp.) 6: 133a
	Biochemical	in vitro chaperone assay with yeast Cdc37	Kimura et al. (1997) Genes & Dev. 11: 1775
	Genetic	Cdc37 required for CKII function and CKII for Cdc37 phosphorylation	Bandhakavi et al. (2003) J. Biol. Chem. 278:2829
Crk1	Genetic & Biochemical	Candida Crk1; 2-hybrid, co-IP, and genetic evidence.	Ni et al. (2004) FEBS Lett. 561: 233
DAPK	Biochemical	co-IP	Citri et al. (2006) J. Biol. Chem. 281: 14361

Table 3.6: (continued...)

Clients	Interaction	Evidence	Reference
EGF receptor mutant	Biochemical	co-IP	Lavictoire et al. (2003) J. Biol. Chem. 278: 5292
Fused	Biochemical	co-IP	Kise et al. (2006) Biochem. Biophys. Res. Com. 351: 78
β-Galactosidase	Biochemical	in vitro chaperone assay with yeast Cdc37	Kimura et al. (1997) Genes & Dev. 11: 1775
Gcn2	Genetic	cdc37 mutant yeast strain grows poorly upon amino acid starvation	Donzé & Picard (1999) Mol. Cell. Biol. 19: 8422
Glucocorticoid receptor	Genetic	Yeast *cdc37-34* has slightly decreased activation of heterologous GR at non-permissive temperature.	Fliss et al. (1997) Mol. Biol. Cell 8: 2501
Harc	Biochemical	Harc and Cdc37 form both homodimers and heterodimers	Roiniotis et al. (2005) Biochem. 44: 6662
Hog1	Genetic & Biochemical	Hog1 function affected by Cdc37 phosphorylation mutant in yeast; co-IP	Hawle et al. (2007) Eukaryot. Cell 6: 521
Hck (src family kinase)	Genetic & Biochemical	Human Cdc37 overexpression partially suppresses tsHck499F phenotype, co-IP, Hck kinase domain necessary for binding	Scholz et al. (2000) Mol. Cell. Biol. 20: 6984
HRI	Biochemical	Human Cdc37 co-IP, maturation intermediates of HRI recruit Cdc37 to Hsp90 heterocomplex	Hartson et al. (2000) Biochem. 39: 7631
	Biochemical	Human Cdc37 associates with nascent HRI co-translationally and persists during maturation and activation, specifically with immature or inactive forms and not with active or repressed forms	Shao et al. (2001) J. Biol. Chem. 276: 206
IKK (IκB kinase)	Biochemical	Human Cdc37 co-IP, in vitro pull downs, direct binding through the kinase domain. Geldanamycin disrupts complex formation and recruitment to receptor.	Chen et al. (2002) Mol. Cell 9: 401; Bouwmeester et al. (2004) Nat. Cell Biol. 6:97
IRAK-1	Biochemical	co-IP	De Nardo et al. (2005) J. Biol. Chem. 280: 9813
JAK1	Biochemical	co-IP	Shang and Tomasi (2006) J. Biol. Chem. 281: 1876
Kin28	Genetic	Synthetic lethal	Valay et al. (1995) J. Mol. Biol. 249: 535
Ksr	Biochemical	part of a larger chaperone complex	Sundaram et al. (1999) Mol. Cell. Biol. 19: 5523
Lck (src family kinase)	Biochemical	Co-IP with human Cdc37	Hartson et al. (2000) Biochem. 39: 7631
Lkb1	Biochemical	Co-IP	Boudeau et al. (2003) Biochem. J. 370: 849; Nony et al. (2003) Oncogene 22: 9165
LRRK2	Biochemical	co-IP	Gloeckner et al. (2006) Hum. Mol. Genet. 15: 223; Wang et al. (2008) J. Neurosci. 28: 3384.
Luciferase (firefly)	Biochemical	in vitro chaperone assay with yeast Cdc37	Kimura et al. (1997) Genes & Dev. 11: 1775
MEKK1/MEKK3	Biochemical	TAP purification	Bouwmeester et al. (2004) Nat. Cell Biol. 6:97
MLK3	Biochemical	co-IP	Zhang et al. (2004) J. Biol. Chem. 279: 19457
MPS1	Genetic	multi-copy suppressor of *mps1-1* & synthetic lethal (with *cdc37-1*)	Schutz et al. (1997) J. Cell Biol. 136: 969
NIK	Biochemical	TAP purification	Bouwmeester et al. (2004) Nat. Cell Biol. 6:97
PDGF receptor α	Biochemical	co-IP	Matei et al. (2007) J. Biol. Chem. 282:445
Pink1	Biochemical	co-IP	Weihofen et al. (2008) Hum. Mol. Genet. 17:602
PKCλ	Biochemical	TAP purification	Brajenovic et al. (2004) J. Biol. Chem. 279: 12804
Raf	Biochemical	Mammalian Cdc37 IP	Perdew et al. (1997) Biochem. 36: 3600
	Biochemical	Mammalian Cdc37 IP, and Hsp90	Stancato et al. (1993) J. Biol. Chem. 268: 21711; Wartmann & Davis (1994) J. Biol. Chem. 269: 6695
	Biochemical	Mammalian Cdc37 co-IP, Cdc37 directly binds Raf kinase domain, separate from TPR domain protein heterocomplexes	Silverstein et al. (1998) J. Biol. Chem. 273: 20090
	Biochemical	Human Cdc37 co-IP, N-terminus of Cdc37 binds Raf1, C-terminus of Cdc37 binds Hsp90, N-terminal 1-163 is dominant negative for Raf activation	Grammatikakis et al. (1999) Mol. Cell. Biol. 19: 1661
RET/PTC1	Biochemical	co-IP	Marsee et al. (2004) J. Biol. Chem. 279: 43990
Reverse Transcriptase (Hepadnavirus)	Biochemical	Human Cdc37 IP and pull downs, specific, direct binding in vitro and in vivo, and Ctermally truncated Cdc37 bound more strongly	Wang et al. (2002) J. Biol. Chem. 277: 24361
Sevenless	Genetic	In Drosophila a genetic link between Sevenless signaling & CDC37, HSP83 & CDC37, p34cdc2 & CDC37	Cutforth & Rubin (1994) Cell 77: 1027
Spc1 (SAPK)	Genetic & Biochemical	*S. pombe*: accumulation and phosphorylation of stress kinase Spc1 is lower in cdc37 mutant; co-IP; Cdc37 requirement may be independent of Hsp90.	Tatebe & Shiozaki (2003) Mol. Cell. Biol. 23: 5132
Slt2	Genetic & Biochemical	Hog1 function affected by Cdc37 phosphorylation mutant in yeast; co-IP, increased for phosphorylated (activated) form of Slt2	Hawle et al. (2007) Eukaryot. Cell 6: 521
v-Src	Biochemical	Mammalian Cdc37 co-IP, with Hsp90	Brugge (1986) Curr. Top. Microbiol. Immunol. 123: 1
	Genetic	Yeast *cdc37-34* & *cdc37-17* mutants have decreased src activity but protein level unaffected	Dey et al. (1996) Mol. Biol. Cell 7: 1405

Table 3.6: (continued...)

Clients	Interaction	Evidence	Reference
	Genetic	Yeast *cdc37-1* had decreased src activity. Genetic link between *HSC82* & *HSP82* with *CDC37*	Kimura et al. (1997) Genes & Dev. 11: 1775
	Biochemical	Mammalian Cdc37 IP	Perdew et al. (1997) Biochem. 36: 3600
Ste11	Genetic & Biochemical	Cdc37 mutant has low Ste11 activity. Pull downs and co-IP	Abbas-Terki et al. (2000) FEBS Lett. 467:111
TAK1	Biochemical	TAP purification	Bouwmeester et al. (2004) Nat. Cell Biol. 6:97
TBK1	Biochemical	TAP purification	Bouwmeester et al. (2004) Nat. Cell Biol. 6:97
Ydj1	Genetic	Synthetic lethal screen	Mort-Bontemps-Soret et al. (2002) Mol. Genet. Genomics 267: 447
ZAP70	Genetic	Rat Cdc37 overexpression restores expression of ZAP70 mutant (protein tyrosine kinase involved in signal transduction throught the T-cell receptor)	Matsuda et al. (1999) J. Biol. Chem. 274: 34515

*Table obtained from www.picard.ch/downloads/Cdc37interactors.pdf

3.5.2 Cdc37 promotes proliferation

Cdc37 was first discovered in a yeast screen for cell division cycle proteins and becomes misappropriated in several cancers. The important role of Cdc37 in supporting Hsp90 activity and kinase function suggests that, like Hsp90, Cdc37 may participate in maintaining malignant phenotype [Pascale et al., 2005; Prince et al., 2005]. Cdc37 was found over expressed in human prostate cancer [Schwarze et al., 2003; Stepanova et al., 2000b]. Cdc37 was also found to be up regulated in hepatocellular carcinoma [Pascale et al., 2005; Prince et al., 2005]. In a genetically engineered mouse model, Cdc37 acted as an oncogene, which was shown by the fact that the targeted over expression in mammary and prostate glands induced proliferative disorders leading to tumours [Stepanova et al., 2000a]. Further, Cdc37 also interacts with oncoproteins c-myc or cyclin D1 and is known to induce the transformation in multiple tissues of mice. Additionally it is known that Hsp90 buffers the oncogenic mutations, and could be also extended the same function to Cdc37. In fact, Cdc37 was first found in mammalian cells in association with Hsp90 and the oncogenic kinase v-Src. Mutation in some of the key kinases involved in the cell signaling allows them to develop 'fast' activity, at the same time also becoming structurally 'loose', and the chaperoning complex of Cdc37-Hsp90 helps them in maintaining this 'fast and loose' state promoting there prolonged existence [Gray et al., 2008].

3.5.3 Cdc37 structure and function

Structurally, the 44.5 kDa protein Cdc37 can be dissected into three domains (Figure 3.8)(MacLean and Picard, 2003; Shao et al., 2003a). Proteolytic fingerprinting studies indicate that it is comprised of an N-terminal domain (residues 1-127 ($Cdc37_N$), 15.5 kDa), a middle domain (residues 147-276

(Cdc37$_M$), 16 kDa), and a C-terminal domain (residues 283-378 (Cdc37$_C$), 10.5 kDa). Additional evidence suggests that the region between 128-147 could comprise an additional domain, as its presence or absence in the deletion constructs affect the kinase-binding and Hsp90-binding. The middle domain Cdc37$_M$ is highly resistant to proteolytic digestion and was found to be the most stable domain of Cdc37 [Zhang et al., 2004]. Several studies [Grammatikakis et al., 1999; Lee et al., 2002; Scholz et al., 2000; Shao et al., 2001, 2003] have shown that Cdc37$_N$ binds to the client proteins (e.g., kinases). A highly conserved region within Cdc37$_N$ contains Ser13 as a unique phosphorylation site for the protein kinase CKII. Phosphorylation seems to be an important mechanism for controlling the binding affinity to the client protein kinases and as a consequence thereof regulating the activity of multiple protein kinases [Miyata and Nishida, 2004]. Recently, it was shown that Hsp90 dependent activation of protein kinases requires chaperone targeted dephosphorylation of Cdc37 by protein phosphatase 5 (PP5) [Vaughan et al., 2008]. This suggests that a cycle of phosphorylation and dephosphorylation of Cdc37 is necessary for kinase-client loading onto Hsp90. The Hsp90 binding site is located in the middle domain of Cdc37, which also comprises a putative dimerization region [Roe et al., 2004]. An extended α-helix connects the Hsp90 binding site to the C-terminal domain, which might also be required for dimer formation [Roe et al., 2004; Siligardi et al., 2002]. The structure of the N-terminal domain is unknown, but sequence analysis predicts a high content of α-helices [Roe et al., 2004].

Cdc37 binds to Hsp90 with a dissociation constant of 1.2 to 3.8 μM depending upon the type of constructs used for the study, viz., human or yeast or Hsp90α or Hsp90β [Roe et al., 2004; Zhang et al., 2004]. The crystal structure of Cdc37$_{MC}$ (middle and C-terminal domain) bound to Hsp90$_N$ (from yeast) is a heterotetrameric complex, comprising of Hsp90 bound to a Cdc37 dimer. The function of Cdc37 to inhibit the ATPase activity of Hsp90 is due to the long side chain of R167 that binds to E33 of Hsp90, thereby inhibiting the ATPase reaction [Roe et al., 2004]. Several structural and biochemical studies have confirmed this interaction is key to the function of this complex. Additional studies also indicate that the full length Cdc37 in its free form exists as a dimer with a dissociation constant of 5-10 μM and 80 μM [Roe et al., 2004]. However there are several contradicting results when taken the results concerning the individual domains [Zhang et al., 2004]. Low resolution cryo-electron microscopy images of the Hsp90-Cdc37-Cdk4 complex show that Cdc37 is a monomer in this interaction [Vaughan et al., 2006]. However, these studies utilized yeast Hsp90 and it was realized that it is necessary to investigate the

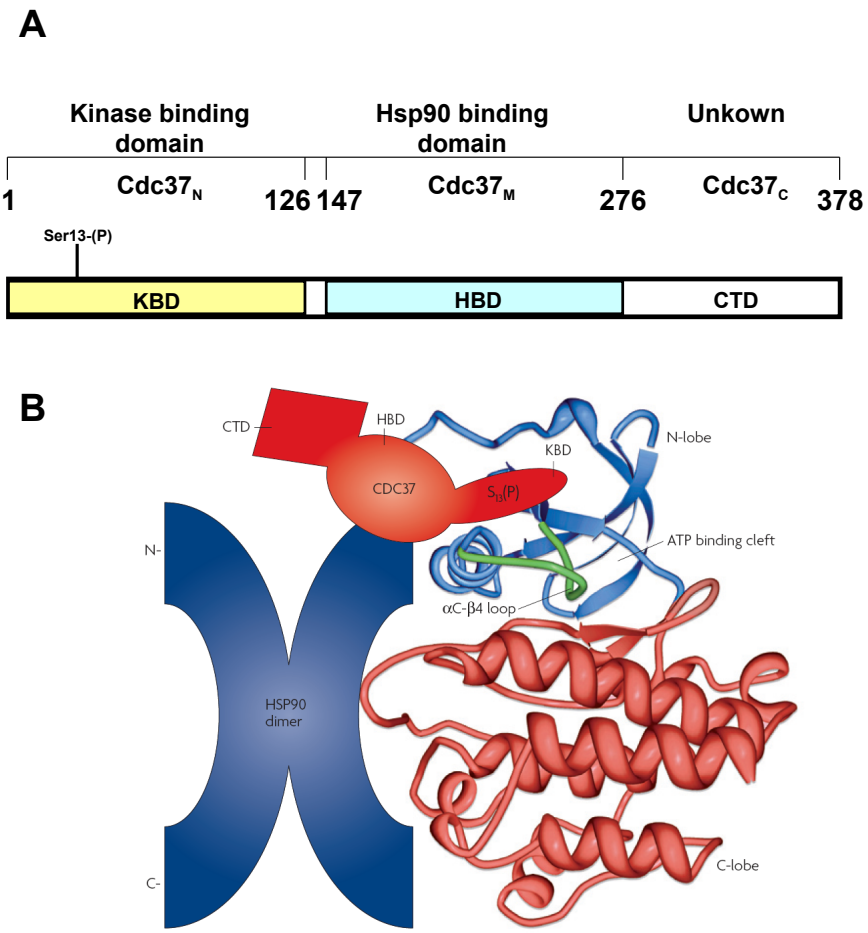


Figure 3.8: Domain architecture of Cdc37. (A) Schematic representation of the three domains of Cdc37. (B) Schematic representation of the present understanding of the ternary complex of Cdc37-Hsp90-kinase. The catalytic domains of protein kinase is represented in ribbon diagram with its N-lobe (blue) and C-lobe (red). The KBD of Cdc37 binds to the N-lobe, HBD binds to N-terminal of Hsp90 and Hsp90 binds both to N and C lobes while interacting with the HBD of Cdc37. Figure 3.8(B) reproduced from Gray et al. [2008].

interaction with human Hsp90 to resolve the differences [Zhang et al., 2004]. In this thesis we determine the structure of the human Hsp90$_N$-Cdc37$_M$ complex by NMR (various NMR methods used to study this protein-protein complex will be discussed in Chapter 5). We also solved the X-ray crystal structure of human Cdc37$_M$ at 1.88 Å in its monomeric form. Based on the available structural and biochemical data the functioning of the Cdc37-Hsp90 chaperone machinery can be summarized as shown in Figure 3.9.

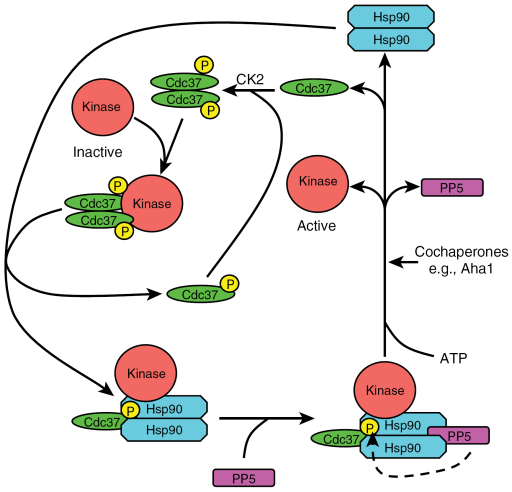


Figure 3.9: Cdc37-Hsp90 chaperone machinery in kinase loading. Cdc37, phosphorylated by CK2, binds to kinases and forms a complex with Hsp90. The PP5 (Ppt1 in yeast) phosphatase binds to the C terminus of Hsp90 and dephosphorylates Cdc37, which subsequently dissociates from the Hsp90 complex. ATP binding and hydrolysis by Hsp90 and interaction with additional cochaperones lead to maturation and release of the kinase. Figure reproduced from Mayer et al. [2009].

3.5.4 Targeting Cdc37 in cancer

Cdc37 is upregulated in cancer cells when compared to the normal tissue. In healthy prostate tissue Cdc37 is barely detectable, whereas it is readily observable in prostate tumor tissue [Stepanova et al., 2000b]. As seen in the earlier section (Cdc37 promotes proliferation), Cdc37 is involved in the maturation of many oncogenic kinase clients and this was even further confirmed by extensive siRNA-mediated silencing of Cdc37. Further, siRNA-mediated silencing of Cdc37 resulted in marked depletion of oncogenic kinases in human colon cancer cell lines[Smith et al., 2009]. Inactivation of Cdc37 may decrease cell survival through multiple pathways, permitting effective inhibition of signaling pathways and also breaking resistance to cytotoxic therapy (Figure 3.10) [Gray et al., 2008; Smith and Workman, 2009]. Therefore targeting of Cdc37 is predicted and proposed to have relatively broad-spectrum anticancer activity due to its diverse set of its client proteins. As compared to inhibition of Hsp90, the client selectivity of Cdc37 co-chaperone is predominantly directed towards protein kinases. Therefore targeting of Cdc37 may be beneficial in kinase-addicted cancers. Additionally, targeting of Cdc37 will be beneficial in prostate cancer, in which androgen receptor(AR) signaling is critical since the AR is an atypical, non-kinase client of Cdc37.

The potential approaches to pharmacologically target Cdc37 activity with small molecules can fall into two categories:

1. **Inhibition of activating phosphorylation of Cdc37 [Smith and Workman, 2009]**- In humans CKII phosphorylates Cdc37 on Ser13 and is known to be essential for modulation of the chaperone cycle. Inhibition of CKII with

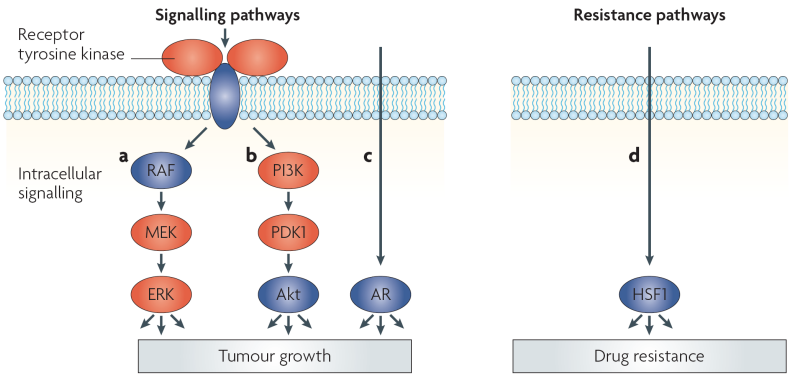


Figure 3.10: Effects of depletion of Cdc37. Cdc37 depletion inhibits signal flux through at least three pathways in prostate carcinoma. These include the (a) extracellular signal-regulated kinase (ERK), (b) Akt and (c) androgen receptor (AR) pathways. Depletion of Cdc37 also inhibits (d), the HSF1-mediated heat shock stress response that can be activated by many anticancer agents. Well-characterized molecular clients for Cdc37 are indicated by blue colouration. MEK, mitogen-activated protein kinase; PDK1, 3-phosphoinositide-dependent kinase 1; PI3K, phosphatidylinositol 3-kinase. Figure reproduced from Gray et al. [2008].

its inhibitor TBB (4,5,6,7-Tetrabromo-2-azabenzimidazole) suppressed Cdc37 phosphorylation and reduced the expression of kinase clients. Thus blocking CKII-mediated activation of Cdc37 could be a possible route of therapeutic intervention. Additionally PP5 is capable of targeted dephosphorylation of Cdc37 within Hsp90 complexes. Over expressing PP5 in HCT116 human colon cancer cells reduced the level of phosphorylated Cdc37, concomitantly decreasing the expression of CRAF client protein kinase expression, demonstrating that the kinase clients are also compromised. Hence modulation of PP5 may offer an additional approach to reduce Cdc37 activity, but this would require more detailed studies as its biochemical interactions are still not well characterized.

2. **Blocking functional protein-protein interactions that Cdc37 is involved [Smith and Workman, 2009]**- Designing small molecule inhibitors for interfering with specific protein-protein interactions is technically more challenging and needs extensive structural characterization and also identifying the 'Hot spots' at the interface between the two proteins. Inhibition of Hsp90 or other client proteins interacting with Cdc37 is another way to inhibit the function of Cdc37. The interaction interface of the protein-protein complex Cdc37-Hsp90, is well characterized structurally and has identified key residues involved in the complex formation. Additionally, in this thesis we also have identified a crucial 'Hot spot' residue which plays a key role in complex formation of

human Hsp90 and human Cdc37 (see Chapter 7). Low molecular weight molecules that interfere with Cdc37 or Hsp90 or disrupt the Hsp90-Cdc37 complex have recently been proposed as a new class of anti-cancer agents [Gray et al., 2008; Smith and Workman, 2009]. Gene based-expression studies have identified the triterpene celastrol, representing a new class of non ATP-competitive inhibitors of Hsp90 [Hieronymus et al., 2006]. Immunoprecipitation in a pancreatic cell line and docking experiments suggested that celastrol exerts its anti-proliferative activity by binding to the N-terminal domain of Hsp90, thereby disrupting the complex between $Hsp90_N$ and Cdc37 [Zhang et al., 2008a]. *In vivo*, celastrol showed significant inhibition of tumor growth in nude mice with prostate or pancreatic cancer [Yang et al., 2006; Zhang et al., 2008a]. In contrast to these studies, in this thesis (see Chapter 8) we elucidated that celastrol, which was thought to bind to Hsp90, in fact binds to Cdc37 and disrupts this protein-protein complex.

4 Small Molecule Inhibitors for Disrupting Protein-Protein Interaction

Protein-protein interactions control and play a key role in biological processes, including cellular growth and differentiation [Alberts, 1998; Gavin et al., 2002]. Any loss of coordination in maintaining these protein-protein interaction becomes a triggering factor for many pathological processes, for example cancer and Alzheimer's disease [Cohen and Prusiner, 1998; Selkoe, 1998]. These protein-protein interactions are very specific in function and hence offer a highly selective target for pharmacological intervention. Thus, inhibition or disruption of these interactions offers a promising novel approach for rational drug design against a wide number of cellular targets. Discovering small molecule drugs to disrupt protein-protein interaction is an enormous challenge [Wells and McClendon, 2007]. In this chapter we will look at some of the examples from the literature, which attempts to disrupt the protein-protein interaction either by using antibodies, peptides, and synthetic or natural small molecules. Of particular interest, we will also discuss about disrupting the Cdc37-Hsp90, kinome co-chaperone-chaperone complex using a recently identified triterpene small molecule celastrol.

4.1 Introduction

Most of the biological processes are not carried out by single proteins, but involves many proteins acting together [Alberts, 1998]. For example the cell cycle-dependent degradation of specific proteins involves a coordinated action of at least ten proteins, called as the anaphase-promoting complex (APC) and involves several other proteins in the downstream events until they reach the 20S proteasome, where

finally the target protein is converted to small peptides. Protein-protein interactions are also the key for carrying out fundamental cellular functions like intracellular signal transduction and also maintaining the cytoskeletal architecture of the cell. In humans, an estimate of roughly 40,000-200,000 protein-protein interactions (human interactome) are thought to exist [Bork et al., 2004]. Only 20,000-30,000 are recorded in the literature for the human interactome. Inappropriate protein-protein interaction can lead to diseases like cancer [White et al., 2008]. Hence, selective small molecule targeting of protein-protein interactions is of great interest for pharmaceutical sciences [Arkin, 2005; Arkin and Wells, 2004; Chène, 2006; Fry and Vassilev, 2005; Sharma et al., 2002; Toogood, 2002; Yin and Hamilton, 2005]. Table 4.1 provides various examples of protein-protein interactions which have been targeted either by using antibodies, peptides, small molecules, or natural products [Loregian and Palù, 2005].

Table 4.1: Overview of protein-protein interactions inhibitors

Target protein-protein interaction	Type of inhibitor	References*
Enzyme subunit interactions		
HSV-1 ribonucleotide reductase (R1/R2)	Peptide	Dutia et al. (1986); Cohen et al. (1986)
	Small molecule	Liuzzi et al. (1994)
HSV-1 DNA polymerase (UL30/UL42)	Peptide	Marsden et al. (1994); Digard et al. (1995)
	Small molecule	Pilger et al. (2004)
HCMV DNA polymerase (UL54/UL44)	Peptide	Loregian et al. (2003)
HIV protease dimerization	Peptide	Zhang et al. (1991)
	Small molecule	Song et al. (2001)
	Natural product	Fan et al. (1998)
HIV reverse transcriptase dimerization	Peptide	Divita et al. (1994)
HIV integrase dimerization	Peptide	Sourgen et al. (1996)
iNOS dimerization	Small molecule	McMillan et al. (2000)
Human gluthatione reductase dimerization	Peptide	Nordhoff et al. (1997)
Ligand/receptor interactions		
Angiotensin II/angiotensin II receptor	Peptide	Moore and Fulton (1984)
	Small molecule	Duncia et al. (1992)
Fibrinogen/integrin GPIIbIIIa	Antibody	Vorchheimer et al. (1999)
	Peptide	McDowell and Gadek (1992)
	Small molecule	McDowell et al. (1994)
Vitronectin/integrin aVb3	Small molecule	Keenan et al. (1998)
MMP2/integrin aVb3	Small molecule	Boger et al. (2001)
VCAM/integrin a4b1	Small molecule	Jackson et al. (1997)
MAdCAM/integrin a4b7	Small molecule	Sircar et al. (2002)
ICAM/LFA-1	Antibody	Marecki and Kirkpatrick (2004)
	Small molecule	Gadek et al. (2002)
IL-1/IL-1 receptor	Peptide	Yanofsky et al. (1996)
	Sc antibody	Chrunyk et al. (2000)
	Small molecule	Sarabu et al. (1997)
IL-2/IL-2 receptor a	Antibody	Ockey and Gadek (2002)

Table continued...

Table 4.1: (continued...)

Target protein-protein interaction	Type of inhibitor	References*
	Small molecule	Tilley et al. (1997)
IL-4/IL-4 receptor a	Peptide	Domingues et al. (1999)
C5a/C5a receptor	Peptide	Finch et al. (1999)
NGF/p75 receptor	Small molecule	Owolabi et al. (1999)
Grb2 SH2/EGF receptor	Small molecule	Fretz et al. (2000)
Grb2 SH2/c-Met	Small molecule	Atabey et al. (2001)
p56lck/T-cell receptor	Small molecule	Proudfoot et al. (2001)
VEGF/KDR receptor	Peptide	Fairbrother et al. (1998)
Rhodopsin/Gt	Peptide	Martin et al. (1996)
Gt a subunit/PDE	Peptide	Skiba et al. (1995)
Rhodopsin/arrestin	Peptide	Krupnick et al. (1994)
Immunophilin protein FKB12/activin receptor R1	Small molecule	Huang and Schreiber (1997)
PI 3-kinase (p85 subunit)/PDGF-b receptor	Peptide	Eaton et al. (1998)
	Small molecule	Eaton et al. (1998)
GRK3/GR	Peptide	Koch et al. (1993)
ZAP-70/T-cell receptor	Peptide	Wange et al. (1995)
PKA/AKAP	Peptide	Carr et al. (1992)
PKC/RACK1	Peptide	Ron and Mochly-Rosen (1994)
IgG Fc/protein A of Staphylococcus aureus	Peptide	DeLano et al. (2000)
HIV gp120/CD4 receptor	Peptide	Ferrer and Harrison (1999)
	Small molecule	Chen et al. (1992)
HIV/RANTES/CCR5	Small molecule	Baba et al. (1999)
erB receptor self-association	Peptide	Berezov et al. (2002)
hER-a self-association	Peptide	Yudt and Koide (2001)
ADR homodimerization	Peptide	Hebert et al. (1996)
Other protein-protein interactions		
Bcl-2/Bak (BH3 domain)	Small molecule	Wang et al. (2000); Degterev et al. (2001)
	Natural product	Tzung et al. (2001)
p53/MDM2	Peptide	Picksley et al. (1994)
	Small molecule	Stoll et al. (2001)
	Natural product	Duncan et al. (2001)
Smac/IAPs	Peptide	Liu et al. (2000)
CaM/ATPase	Small molecule	Levin and Weiss (1980)
CaM/smMLCK	Small molecule	Orner et al. (2001)
CaM/CaM-dependent kinases	Peptide	Nevalainen et al. (1997)
Tubulin-a/b polymerization	Small molecule	Haggarty et al. (2000)
	Natural product	Usui et al. (1998)
Hsp90/p23	Small molecule	Chiosis et al. (2001)
a1B/b3 subunits of human N-type calcium channel	Small molecule	Young et al. (1998)
IGF-1/IGFBP-1	Peptide	Lowman et al. (1998)
Stat3 dimerization	Peptide	Ren et al. (2003)
HIV p24 polymerization	Peptide	Hilpert et al. (1999)

aADR, β2-adrenergic receptor; AKAP, A-kinase anchoring protein; Bcl-2, B-cell leukemia/lymphoma 2; BH3, Bcl-2 homology domain 3; CaM, calmoduline; CCR5, CCchemokine receptor 5; c-Met, hepatocyte growth factor receptor; EGF, epidermal growth factor; GPIIbIIIa, glycoprotein IIbIIIa; GR, G-protein-coupled receptor; GRK3, G-protein-coupled receptor kinase 3; Grb2, growth factor receptor bound protein 2; Gt, G protein transducin; HCMV, human cytomegalovirus; hER-α, human estrogen receptor-α; HIV, human immunodeficiency virus; Hsp90, heat shock protein 90; HSV-1, herpes simplex virus type 1; IAPs, inhibitor of apoptosis proteins; ICAM-1, intercellular cell adhesion molecule 1; IGF-1, insuline-like growth factor 1; IGFBP-1, insuline-like growth factor binding protein 1; IL, interleukin; iNOS, inducible nitric oxide synthase; KDR, kinase insert domain containing receptor; LFA-1, lymphocyte functional antigen 1; MAdCAM, mucosal addressin cell adhesion molecule; MMP2, matrix metalloproteinase 2; NGF, nerve growth factor; PDE, phosphodiesterase; PDGF-β, platelet-derived growth factor-b; PI, phosphatidyl-inositol; PKA, cAMP-dependent serine/threonine protein kinase A; PKC, protein kinase C; RACK1, receptor for activated C-kinase 1; sc antibody, single chain antibody; SH2, Src homology domain 2; smMLCK, small muscle myosin light chain kinase; Stat3, signal transduction and activation of transcription 3; VCAM, vascular cellular adhesion molecule; VEGF, vascular endothelial growth factor. *For back references please refer Loregian and Palù [2005]

4.1.1 Major challenges and approaches used for targeting protein-protein interactions

Targeting protein-protein interactions as a strategy for preventing diseases is attractive, but it also comes at the cost of major difficulties to be overcome before an inhibitor could be designed. The difficulties derive mainly because of either one or more of the following nature of protein-protein interactions [Arkin, 2005; Arkin and Wells, 2004; Cochran, 2000; Loregian and Palù, 2005; Wells and McClendon, 2007],

- The interaction interface might be very large and flat
- Very few residues at the interface might contribute to the binding within the contact surfaces
- A single point mutation in one of the subunit of the protein-protein interface can lead to the disruption of the subunit interaction or reduce its binding affinity significantly
- Affinity of protein-protein interactions

The nature of protein-protein interfaces vary widely and not all interactions are amenable for drug discovery as others do.

Several approaches have been tried to target protein-protein interactions viz.,

- Natural occurring small molecules. Example: Taxol [Downing, 2000] and rapamycin [Choi et al., 1996]
- Small molecule substrates of enzymes as starting template for designing antagonists. Examples include several enzymes
- Mapping the epitope of the proteins onto a small peptide or peptidomimetic. Example: peptide targeting integrins GPIIbIIIa, αvβ3 and α4β1 [Arkin and Wells, 2004]
- Random screening. Example: Vinca alkaloids against tubulin polymerization
- Identification of 'Hotspots' (centralized regions of residues that are crucial for the affinity of the interaction) by various biophysical techniques, at the binding interface of protein-protein interface aids in efficient design of drugs. Example: Ro26-4550 an inhibitor of a cytokine-receptor interaction [Arkin and Wells, 2004]

4.1.2 Small molecule inhibitor of Hsp90-Cdc37 complex

In chapter 3, subsection 3.5.4, we have seen the importance of targeting the protein-protein complex Hsp90-Cdc37. This protein-protein complex forms with a K_D = 1.2 μM, and is considered to mediate carcinogenesis by stabilizing a variety of different oncogenic kinases in malignant cells. Low molecular weight molecules that interfere with Cdc37 or Hsp90 or disrupt the Hsp90-Cdc37 complex have recently been proposed as a new class of anti-cancer agents [Gray et al., 2008; Smith and Workman, 2009]. Gene based-expression studies have identified the triterpene celastrol, representing a new class of non ATP-competitive inhibitors of Hsp90 [Hieronymus et al., 2006]. Immunoprecipitation in a pancreatic cell line and docking experiments suggested that celastrol exerts its anti-proliferative activity by binding to the N-terminal domain of Hsp90 ($Hsp90_N$), thereby disrupting the complex between $Hsp90_N$ and Cdc37 [Zhang et al., 2008a]. In vivo, celastrol showed significant inhibition of tumor growth in nude mice with prostate or pancreatic cancer [Yang et al., 2006; Zhang et al., 2008a].

However, in this thesis we established by NMR that celastrol disrupts the Hsp90-Cdc37, by binding to Cdc37 (Figure 4.1), in contrast to the published reports (see Chapter 8). Celastrol inactivates Cdc37 by a thiol-mediated mechanism (Figure 4.2).

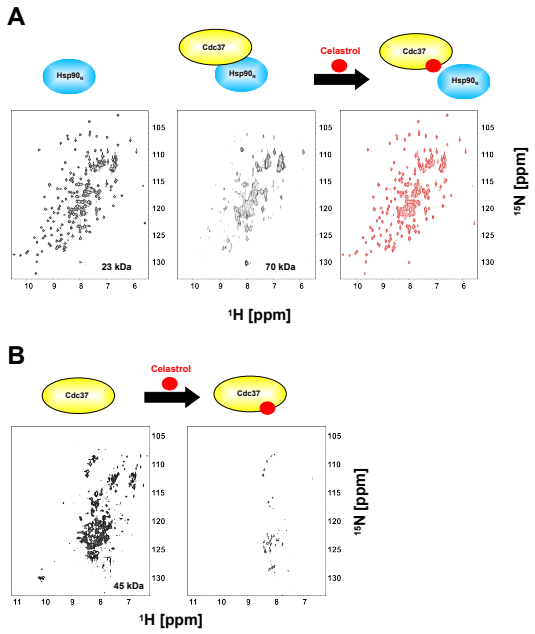


Figure 4.1: Binding of celastrol to Cdc37.
A) (left) ^{1}H,^{15}N HSQC spectrum of Hsp90$_N$. (middle) ^{1}H,^{15}N HSQC spectrum of Hsp90$_N$-Cdc37 complex. (right) Upon addition of celastrol (1:7 of Hsp90$_N$-Cdc37:celastrol), the complex dissociates and the spectrum of free Hsp90$_N$ is retained indicating that celastrol does not bind to Hsp90$_N$. B) (left) ^{1}H,^{15}N HSQC spectrum of full-length Cdc37 (45 kDa). (right) The cross-peaks disappear upon addition of celastrol (1:7 of Cdc37:celastrol), indicating a complex formation with Cdc37.

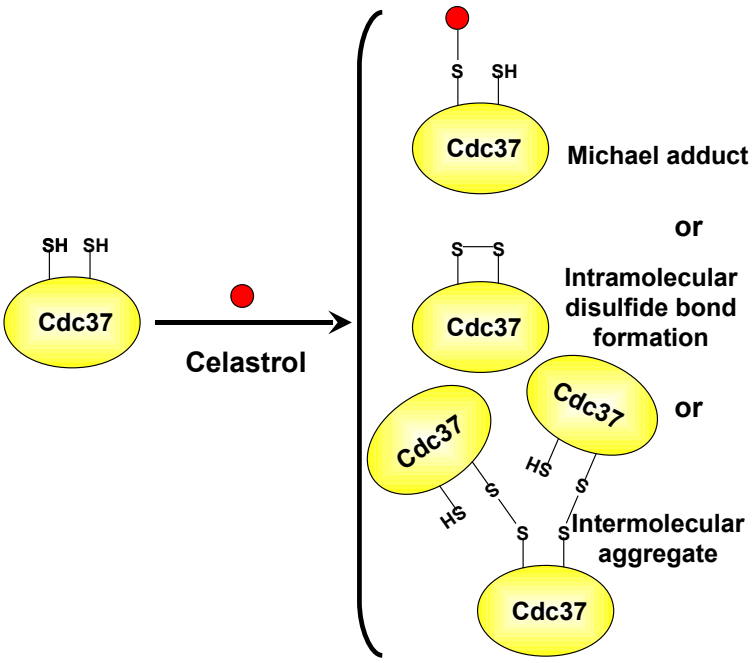


Figure 4.2: Mechanism of inactivation of Cdc37 by celastrol. Celastrol inactivates Cdc37 either by the formation of a covalent adduct(s) with thiol(s) of Cdc37, by formation of an intramolecular disulfide bond or by intermolecular aggregates.

5 NMR Methods to Study Protein-Protein interactions

5.1 Introduction

Characterizing protein-protein interaction is vital for understanding the molecular mechanism of cellular processes [Alberts, 1998]. Several experimental techniques have been used to characterize protein interactions. Table 5.1 lists some of the genetic, biochemical and biophysical techniques used to analyze protein interactions. Some of these techniques are used for screening a large number of protein interactions in cell, viz., Yeast-2-hybrid system (Y2H), tandem affinity purification (TAP) etc., and other methods mainly are used to characterizing the biochemical, and physico-chemical properties of interaction.

Table 5.1: Different methods used to characterize protein-protein interactions

Method	High-throughput approach	Living cell assay	Type of interactions	Type of characterization
Y2H	+	In vivo	Physical interaction (binary)	Identification
Affinity purification-MS	+	In vitro	Physical interaction (complex)	Identification
DNA microarrays/gene coexpression	+	In vitro	Functional association	Identification
Protein microarrays	+	In vitro	Physical interaction (complex)	Identification
Synthetic lethality	+	In vitro	Functional association	Identification
Phage display	+	In vitro	Physical interaction (complex)	Identification
X-ray crystallography	-	In vitro	Physical interaction (complex)	Structural and biological characterization

Table 5.1: (continued...)

Method	High-throughput approach	Living cell assay	Type of interactions	Type of characterization
NMR spectroscopy	-	In vitro	Physical interaction (complex)	Kinetic, dynamic, structural and biological characterization
Fluorescence resonance energy transfer	-	In vitro	Physical interaction (binary)	Dynamic and Biological characterization
Surface plasmon resonance	-	In vitro	Physical interaction (complex)	Kinetic and dynamic characterization
Atomic force microscopy	-	In vitro	Physical interaction (binary)	Mechanical and dynamic characterization
Electron microscopy	-	In vitro	Physical interaction (complex)	Structural and biological characterization

In this chapter, we will discuss more about Nuclear Magnetic Resonance(NMR) spectroscopy, which is the major technique used in this thesis. We will specially focus its application towards the study of protein-protein interaction.

5.2 Nuclear Magnetic Resonance (NMR) methods to study proteins and protein complexes

5.2.1 Introduction

NMR spectroscopy is an experimental tool developed independently by Felix Bloch and Edward Purcell in the year 1945, who were primarily interested in exploring fundamental properties of matter. Since then in sixty years its applications has grown enormously than ever they could have imagined. One such area of application is structural biology, where in it is used to elucidate the structure and function of biomacromolecules in the fascinating world of cellular organization. It has become the only technique with the ability of studying kinetics, dynamics, and structure determination at atomic resolution in solution and near to physiological condition. Figure 5.1 shows timescale of dynamic processes of proteins and the various experimental techniques that can detect on each timescale, with special emphasis on NMR parameters. In contrast to the three-dimensional static structures of proteins determined by X-ray crystallography, NMR adds the fourth dimension 'time' and thereby reveals the dynamic personality of these proteins, which becomes fundamental for describing cellular function of proteins.

NMR phenomena is a property of all the atomic nuclei called the 'spin' and the most

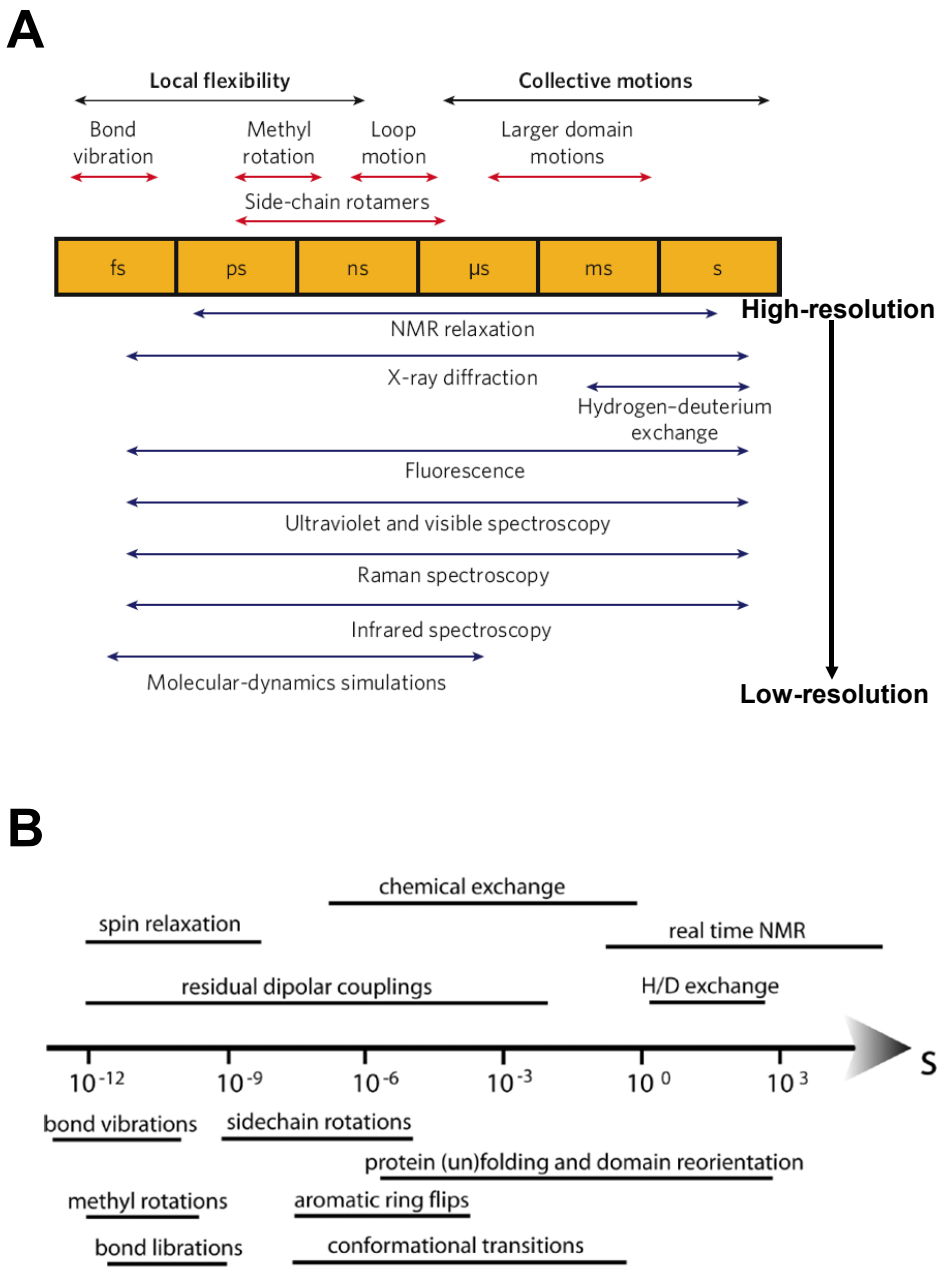


Figure 5.1: Protein dynamics and methods to study. (A) Timescale of dynamic processes in proteins and the experimental methods that can detect fluctuations on each timescale (modified from Henzler-Wildman and Kern [2007]). (B) The different processes amenable to NMR spectroscopy are indicated above the time arrow, below are the typical time windows for different molecular motions and events.

commonly studied nuclei are the spin $\frac{1}{2}$ particles, ^{1}H, ^{13}C, ^{15}N and ^{31}P. Proteins are rich in protons and this property has been utilized and the first 1D NMR spectra of a protein (ribonuclease) was recorded in 1957 [Jardetzky and Jardetzky, 1957]. However, as proteins contain a large number of atoms, and hence it is hard to expect that the resonance lines of individual atoms could be resolved by a one dimensional spectra. Hence the current technological advancement in NMR allows us to extend to multidimensional NMR experiments. Multidimensional experiments are not limited to a single type of atoms such as protons (^{1}H-^{1}H COSY, ^{1}H-^{1}H TOCSY and ^{1}H-^{1}H NOESY). By isotopically enriching the proteins with ^{15}N and ^{13}C, an additional frequency domain can be added, thereby increasing the spectral resolution and the information gained from multidimensional experiments. Important is the heteronuclear single quantum coherence (HSQC) experiment that correlates the resonance frequency of a proton with that of the directly bonded hetero nucleus, example, the amide proton resonance with amide nitrogen resonance. Table 5.2 summarizes a selection of important NMR experiments.

Table 5.2: Important nD experiments

Experiment	Correlated resonances	References
COSY	Directly J-coupled spins	[Aue et al., 1976]
TOCSY	All spins of a J-coupled spin system	[Braunschweiler and Ernst, 1983]
NOESY	All spins in close neighborhood ($<$ 0.5 nm)	[Jeener et al., 1979]
HSQC	Protons with directly bound to ^{15}N or ^{13}C, 'fingerprint'	[Bax et al., 1990b; Bodenhausen and Ruben, 1980; Müller, 1979; Norwood et al., 1990]
HNCA	H_i^N with N_i^H and C_i^α and C_{i-1}^α	[Farmer II et al., 1992; Grzesiek and Bax, 1992c; Kay et al., 1990]
HNCO	H_i^N with N_i^H and $C_{i-1}^{'}$	[Grzesiek and Bax, 1992c; Kay et al., 1990; Muhandiram and Kay, 1994]
HN(CO)CA	H_i^N with N_i^H and C_{i-1}^α	[Bax and Ikura, 1991; Grzesiek and Bax, 1992a]
HCCH-TOCSY	Hα with *via* ^{13}C coupled protons of the same amino acid	[Bax et al., 1990a; Olejniczak et al., 1992]
CBCA(CO)NH	C_{i-1}^α and $C_{i-1}^\beta$with N_i^H and H_i^N	[Grzesiek and Bax, 1992a, 1993]
CBCANH	C_i^α, C_i^β, C_{i-1}^α and C_{i-1}^β with N_i^H and H_i^N	[Grzesiek and Bax, 1992b]
HN(CO)CACB	H_i^N with N_i^H and C_{i-1}^α and C_{i-1}^β	[Yamazaki et al., 1994]
HNCACB	H_i^N with N_i^H and C_i^α, C_i^β, C_{i-1}^α and C_{i-1}^β	[Muhandiram and Kay, 1994; Wittekind and Mueller, 1993]

However, the major disadvantage of NMR to study proteins and protein-protein complexes is its size limitation. In the next subsection 5.2.2 we will see how one typically tries to overcome this problem by using a combination of advanced NMR experiments and protein properties.

5.2.2 Methods to study large proteins: use of advanced NMR methods and optimization of protein domains to study protein-protein interactions

Studying high molecular weight proteins and macromolecular complexes by NMR, is limited by number of factors, viz.,

1. high complexity of the spectra due to extensive signal overlap
2. decreased sensitivity due to low concentrations of the sample
3. rapid transverse relaxation leads to even low sensitivity and significant line broadening of the signals

5.2.3 NMR methods

For large molecules, the relaxation rate is $1/T_2$ and thus the linewidth $\Delta\nu_{1/2} = 1/(\pi T_2)$ is proportional to the rotational correlation time τ_c. Since the rotational correlation time, τ_c is proportional to the molecular weight, the T_2 is inversely proportional to the molecular weight(MW)(Equation 5.1).

$$T_2 = \frac{1}{MW} \tag{5.1}$$

As the size of the molecule increases, the T_2 becomes shorter. In larger molecules, it becomes increasingly more difficult to transfer magnetization between spins because of the relaxation losses that are incurred during the polarization transfer periods. There are two methods by which the T_2 of the spins can be lengthened. The first involves changing the rotational correlation time of the sample by changing the temperature, since τ_c for a spherically rigid molecule will follow the Stokes-Einstein equation (Equation 5.2).

$$\tau_c = \frac{V\eta}{\kappa T} \tag{5.2}$$

where V the volume of the protein, η the viscosity, and T the absolute temperature. The second method utilizes relaxation interference between dipolar coupling and the chemical shift anisotropy. Considerable effort in this direction has been made to overcome this limitation using deuteration, creative labeling schemes and advanced pulse sequences such as TROSY [Pervushin et al., 1997] (transverse relaxation optimized spectroscopy), CRIPT [Riek et al., 1999] (cross-correlated relaxation-induced polarization transfer) and CRINEPT [Riek et al., 1999] (cross-correlated relaxation-enhanced polarization transfer). Thus it has become possible to record well resolved spectra and obtain some structural details for systems up to 900 kDa [Fiaux et al., 2002; Horst et al., 2005; Kreishman-Deitrick et al., 2003]. Methyl labeling strategies has enabled resonance assignment and structure calculation of an 82 kDa monomeric protein [Tugarinov et al., 2005; Tugarinov and Kay, 2004; Tugarinov et al., 2002].

5.2.4 Optimization of protein domains for NMR studies

Apart from the above described NMR methods, the most commonly used method to overcome high-molecular weight problem is the " divide-and-conquer" approach, where constituent domains of a protein are individually cloned, expressed, and purified. Theoretically, this strategy reduces the τ_c by decreasing the molecular weight, and there by increases the sharpness of the resonance signals.

Some of the methods used to identify the domain boundaries in proteins include [Card and Gardner, 2005],

1. Computational methods
    - Domain prediction methods
    - Secondary structure prediction methods
2. Experimental methods
    - Parallel fragment cloning and expression
    - Choice of fusion protein
    - *E. coli* cell lysate screening by NMR

Computational methods: Robust algorithms has been designed to identify and annotate protein domains and secondary structural elements to locate independently folded, stable protein constructs that can be cloned and

expressed. Domain prediction methods utilizes the Hidden Markov Models [Krogh et al., 1994] (HMM), wherein multiple alignments are based on the assumption that structurally similar proteins could still vary largely in sequence, but there should be conserved elements within the primary sequence that defines the folding, structure and function of the target. Pfam and SMART libraries (a Simple Modular Architecture Research Tool) are such tools based on this algorithm and are widely used. SMART also offers additional advantage of identifying compositionally biased regions such as transmembrane segments, which can assist in the rational design of protein fragments with less aggregation properties for concentrations required for NMR. Secondary structure prediction methods such as PHD, PREDATOR and JPred also assist in identifying secondary structures and domain boundaries.

Experimental methods: Limited proteolytic digestion is one of the common methods used to identify the stable domains in proteins [Cohen et al., 1995]. In this thesis we have utilized this strategy for optimizing the domains of 45 kDa protein Cdc37 for NMR study (Figure 5.2). Additionally one could also use parallel expression approach, wherein multiple N- and C-termini flanking domains identified previously by computational methods are cloned with fusion proteins (example protein GB1) to enhance solubility and then screened the expressed protein by recording a 2D ^{1}H,^{15}N HSQC, whose foldedness can be easily assessed by the dispersion of the backbone amide proton chemical shifts. Additionally, for proteins less than 20 kDa, direct measuring of 2D ^{1}H,^{15}N HSQC of the *E. coli* cell lysate from cultures over expressing proteins is very useful. This method has been used for domain optimization of N-terminal PAS domain of human PAS kinase [Amezcua et al., 2002].

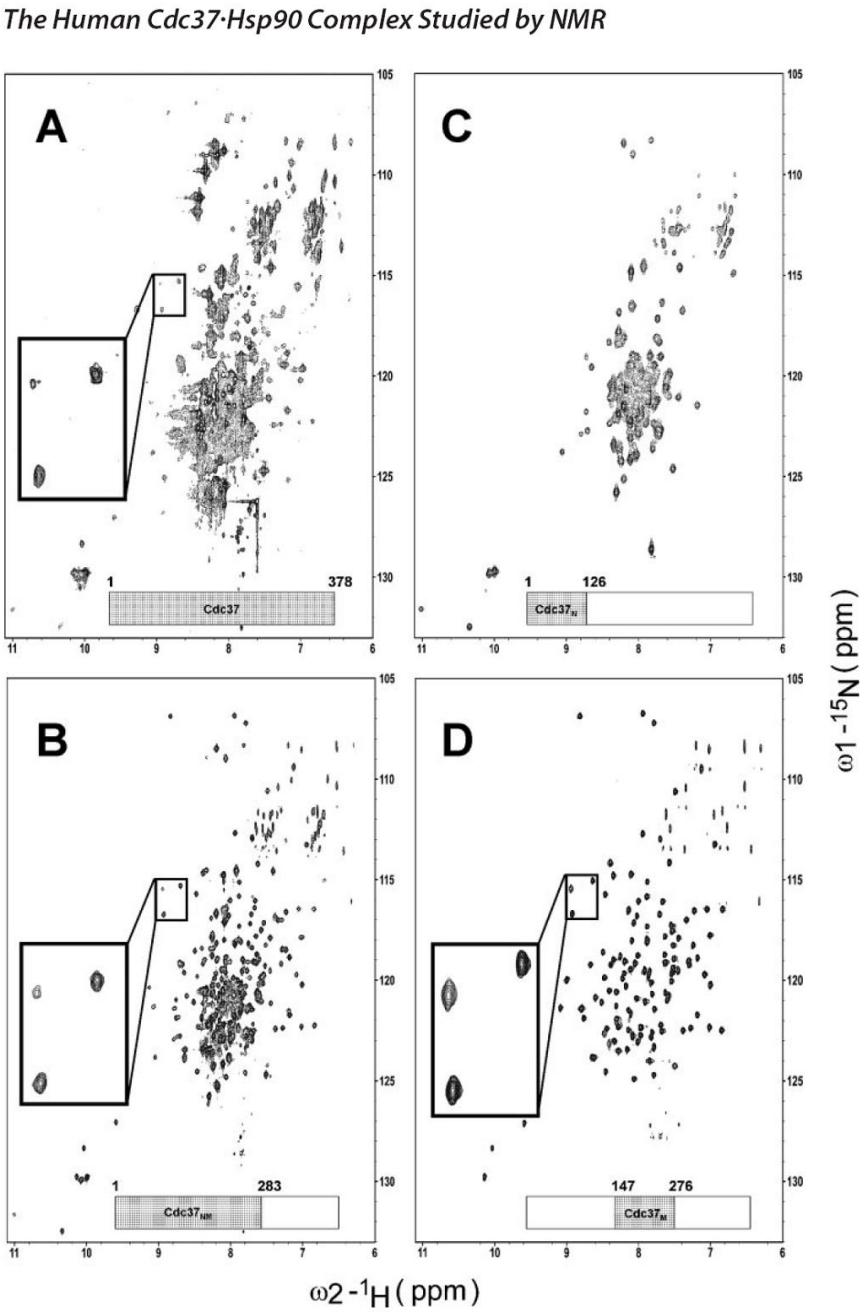


Figure 5.2: ^{1}H,^{15}N TROSY spectra (950 MHz) of different Cdc37 constructs. The boundaries of each construct (eg: (A) full length, (B) N-terminal and Middle domain, (C) N-terminal domain, (D) Middle domain) are schematically indicated below each spectrum. The inlay shows three peaks stemming from the middle domain (Cdc37$_M$), showing the same fold for this domain in all constructs. The spectra were acquired at 298 K in the same buffer solution [50 mM Hepes buffer (pH 7.5), 100 mM NaCl, 1 mM DTT and 10% D2O].

5.2.5 NMR based methods for the study of protein-protein interaction [Zuiderweg, 2002]

NMR provides the most informative but also is the most difficult method for studying large protein-protein complexes and its structure determination. In most of the applications it is sufficient just to map the protein-protein interface (or the interaction site) and then use this information to dock the two proteins or protein-ligand. Some of the important methods used for mapping the protein-protein interface are,

1. Nuclear Overhauser Effect (NOE)
2. Chemical Shift Perturbation (CSP)
3. Cross-Saturation Transfer (CST)
4. Mapping with Dynamics
5. Mapping with Amide-Proton Exchange (H-D Exchange)
6. Mapping with Paramagnetics
7. Mapping with site-specific Spin-labeling
8. Relative positions of proteins using Residual Dipolar Couplings (RDC)

5.2.5.1 Nuclear Overhauser Effect (NOE)

Mapping of biomacromolecular interactions by using NOE method provides an unambiguous information. NOE measures the interproton distances with a distance proportionality of r^{-6}. Using many NOE derived distance constraints between two interacting partners and in addition to several thousands intramolecular NOEs, one can arrive at a three-dimensional structure of the complex. This method is specially useful to probe tightly binding partners ($k_D = 10\ \mu$M). Isotope edited NOE [Fesik et al., 1988] (half-filter) is the most widely used technique to map protein-protein interface. This requires that the two interacting macromolecules have different isotope labeling patterns. For example, one partner has no labeling (^{1}H, ^{12}C, and ^{14}N) while the other is labeled with stable isotopes, example ^{1}H, ^{12}C, ^{15}N or ^{1}H, ^{13}C, ^{14}N or sometimes even ^{1}H, ^{2}H, ^{13}C, and ^{15}N. The 40 kDa protein-protein complex EIN-HPr (Enzyme I N-terminus-Histidine containing phospho carrier protein) was solved exclusively by this method [Garrett et al., 1999]. Several examples of protein-DNA exist as it is a very suitable method for using labeled protein and

unlabeled nucleic acids. Examples in this class include the complex of the lac repressor protein with lac operator DNA [Spronk et al., 1999].

5.2.5.2 Chemical Shift Perturbation (CSP)

CSP is the widely used NMR method to map the protein interfaces. The chemical shift (resonance frequency) of a given spin in the protein reflects the chemical environment and responds very sensitively to any small changes structurally. In a nutshell, the ^{1}H,^{15}N HSQC spectrum of a protein is monitored when the unlabeled interaction partner is titrated in (NMR titration), and the perturbations of the chemical shifts are recorded. The quantification of chemical shift (δ) is made by calculating the length of the vector that connects the two-dimensional end points, normalized by typical shift ranges [Qin et al., 2001](Equation 5.3).

$$\Delta_{av} = [(\Delta\delta_{HN}^2 + \Delta\delta_N^2/25)/2]^{1/2} \tag{5.3}$$

The maximum shift change resonance also provides estimates for the lower limit of the dissociation rate constant. Some of the examples for which mapping methodology was used include complexes between HPr-eEnzyme-II [Emerson et al., 1995; Garrett et al., 1997], GB1-Fc [Otting, 1993], cRAF-Ras-GMPPNP [Foster et al., 1998], 1Nef-CD4 complex [Grzesiek et al., 1996], etc. Generally, one cannot use chemical shift to know what exactly happens at the interface on a per atom basis, but it gives a picture of the local and global changes induced as a result of complex formation. In other terms, they provide connectivity of putative interaction site. Figure 5.3 shows the correlation of CSPs versus residues in Cdc37$_M$ upon addition of Hsp90$_N$ and vice-versa.

NMR titration experiments in addition to providing the information for mapping of the interface, it also gives a good estimation of the affinity, stoichiometry, and specificity of binding as well as the kinetics of binding. A key factor in the use of NMR for measuring dissociation constants is its sensitivity to the rate of “chemical exchange”.

Chemical exchange and NMR: In NMR, “chemical exchange” refers to any process in which a nucleus exchanges between two or more environments in which its NMR parameters (chemical shift, scalar or dipolar coupling, relaxation) differ. The effect of this exchange process on the appearance of the NMR spectrum depends on the

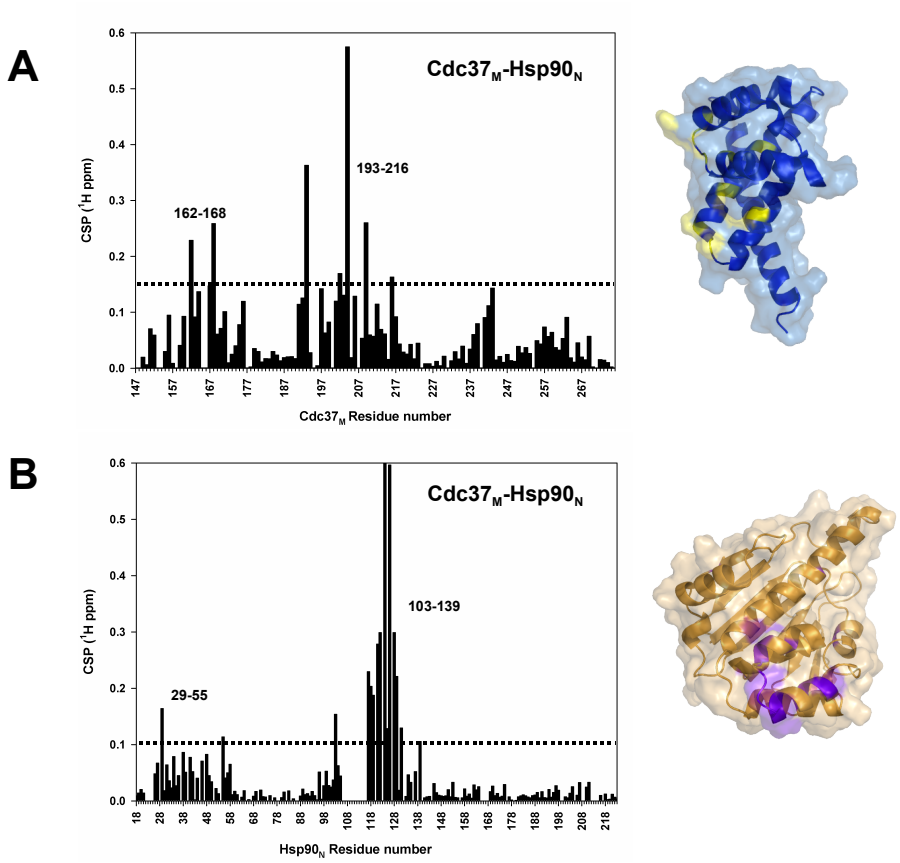


Figure 5.3: Titration of Cdc37$_M$ with Hsp90$_N$ and vice-versa. (A) Correlation of CSPs versus residues in Cdc37$_M$ upon addition of Hsp90$_N$ (left) and surface representation of the interaction sites of Cdc37$_M$ (yellow) for Hsp90$_N$ (> 0.15 ppm). (B) Correlation of CSPs versus residues in Hsp90$_N$ upon addition of Cdc37$_M$ (left) and surface representation of the interaction sites of Hsp90$_N$ (purple) for Cdc37$_M$ (> 0.1 ppm). For the CSPs, amide ^{1}H and ^{15}N resonance shifts have been mapped and combined as Euclidian distances between peak maxima taking into account the gyromagnetic ratio of proton and nitrogen (in ^{1}H ppm). Missing bars indicate prolines or amide resonances that have not been assigned.

rate of exchange relative to the magnitude of the difference in NMR parameters between the two states.

The kinetics of exchange reaction are defined by the following scheme:

$$A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} B \tag{5.4}$$

The kinetic rate constant for the conversion of A to B is k_1 and the rate constant for the reverse reaction is k_2 (Equation 5.4). To characterize the different time scales of exchange it is useful to define an apparent exchange rate, $k_{ex} = k_1 + k_2$, and a frequency difference between the two states, $\Delta\omega = \omega$A - ωB. Table 5.3

summarizes the effects of exchange on the properties of the NMR spectrum. The left columns describe the exchange in terms of the relationship between the frequency separation ($\Delta\nu = 1/2\pi(\omega$A - ωB)) and the apparent exchange rate constant (k_{ex}= $k_1 + k_2$). The remaining columns describe the effect of exchange on the spectrum and the appropriate experimental technique to characterize the exchange rate. A more comprehensive description of the methodologies that can be used to measure chemical exchange can be found in Palmer et al. [2001] Note that the response of the system to chemical exchange depends on the ratio of the exchange rate to the frequency difference of the spins in each environment, i.e. $k_{ex}/\Delta\omega$. Figure 5.4 shows an example of effect of ligand binding on the NMR lineshapes.

Table 5.3: Summary of the effects of exchange on the properties of the NMR spectrum

Chemical exchange type	Exchange rate	α*	Observed spectrum	Experimental technique
Very slow	$k_{ex} << \Delta\nu$	0	Two resonances	Exchange spectroscopy
Slow	$k_{ex} < \Delta\nu$	<1	Two broadened resonances	Line width measurement
Intermediate	$k_{ex} \approx \Delta\nu$	1	Complex line shape	Line shape analysis
Fast	$k_{ex} > \Delta\nu$	>1	Single broadened resonance	Spin-echo R(τ_{cp})
Very fast	$k_{ex} >> \Delta\nu$	2	Single resonance	Spin-lock R(ω_e)

*α is a dimensionless parameter as defined by Millet et al. [2000]

In Figure 5.4(B), one observes that some peaks disappear and reappear at a new chemical shift, and other peaks just shifting in position, would indicate that some peaks are undergoing exchange on a slow to intermediate timescale and some on a fast time scale. The apparently disparate behavior of the peaks arises from the dependence of the line shape on the chemical shift difference between the bound and unbound states. Larger chemical shift differences lead to the appearance of new peak concomitantly with loss of intensity at the original frequency and smaller chemical shift differences give rise to a gradual shifting and broadening of the peak.

5.2.5.3 Cross-Saturation Transfer (CST)

CST [Takahashi et al., 2000] is an highly relevant quick interface mapping technique. Cross-saturation is governed by the same physical process as that of NOE experiment. The donating partner protein is not labeled, while the observed

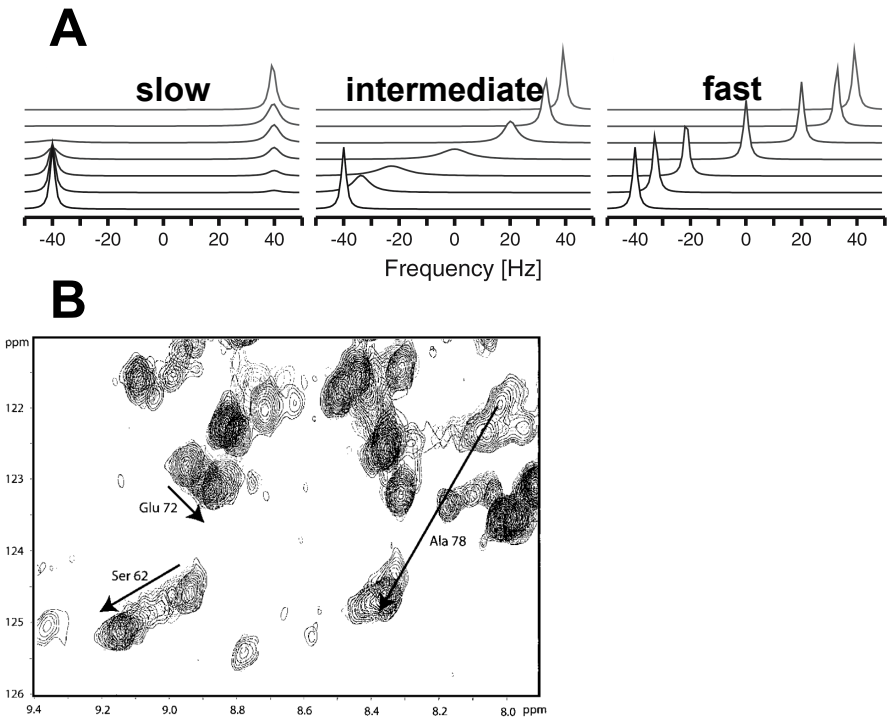


Figure 5.4: Effects of ligand binding on NMR lineshapes. (A)This figure illustrates the effect of slow (left), intermediate (middle), or fast (right) exchange on the spectrum of a resonance, whose frequency is changed as a result of ligand binding.(B) Expanded region of the overlaid ^{1}H, ^{15}N HSQC spectra from a titration of ^{15}N-labeled S100B with CapZ peptide [Kilby et al., 1997]. The protein concentration was 1 mM, and the spectra shown are for peptide concentrations of 0, 0.25, 0.5, 0.75, 1.0, 1.5, 2.5 and 3.0 mM. Chemical shift changes for the cross peaks of Ser62, Glu72 and Ala78 are indicated.

protein is perdeuterated and ^{15}N labeled, but its amide deuterons are exchanged back to protons (Figure 5.5).

The NMR experiment starts with a steady state saturation of, exclusively, all aliphatic proton resonances of the donating partner. Cross-saturation carries the saturation from the donor to the acceptor protein amide protons, where it is detected using ^{1}H,^{15}N HSQC or, for larger proteins, ^{1}H,^{15}N TROSY. Those acceptor cross-peaks that change in intensity upon donor saturation are very likely to be close to the intermolecular interface. The labeling can be reversed to obtain the other interface. This method has been successfully used for mapping the interface between Mbp1(Mlu1 cell cycle binding protein 1)-DNA complex [Lane et al., 2001], Cytochrome b5-Cytochrome c Complex [Deep et al., 2005] etc. In this thesis, we have used this experiment to map the binding interface of $Cdc37_M$-$Hsp90_N$ complex (Figure 5.6).

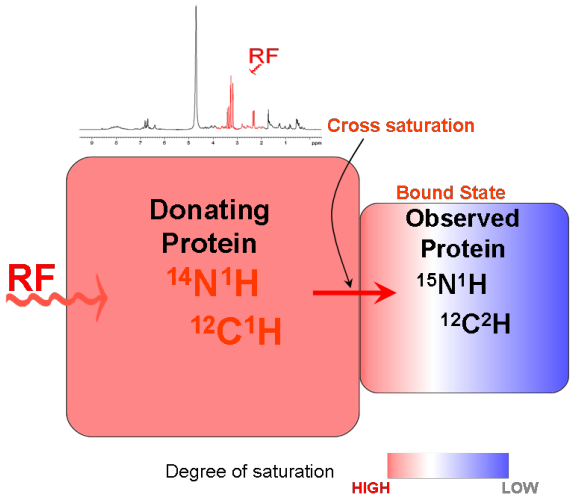


Figure 5.5: Principle of the cross-saturation transfer NMR experiment.

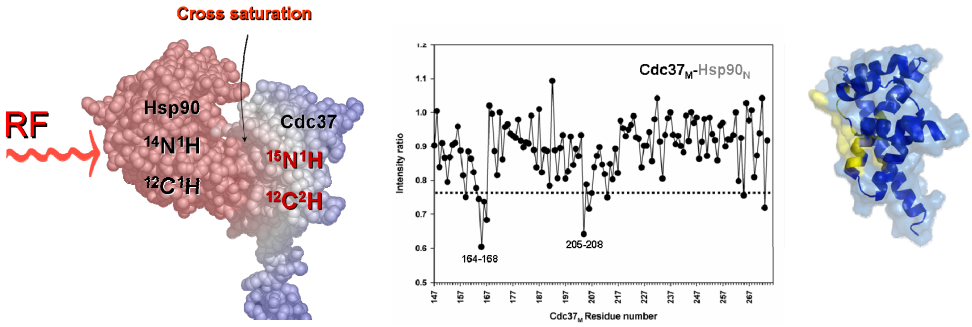


Figure 5.6: Cross-saturation transfer method used for $Cdc37_M$-$Hsp90_N$ complex. Plot of the intensity ratios of the crosspeaks in the cross-saturation transfer experiment. The intensity ratios of the backbone amide resonance of $Cdc37_M$ in complex with $Hsp90_N$ with irradiation to those of without irradiation measured in 10 % H_2O / 90 % D_2O. Surface representation of $Cdc37_M$, where residues with intensity ratio less than 0.78 are colored in yellow.

For weakly binding proteins another variation of the above experiment is transferred cross-saturation experiment (TCS) [Nakanishi et al., 2002]. This method has been successfully used in mapping the epitopes of ligand proteins for binding to a membrane protein (agitoxin2 (AgTx2) to the KcsA K+ channel) [Takeuchi et al., 2003], a membrane (mastoparan bound to the lipid bilayer) [Nakamura et al., 2005] and the heterogenous system of the extracellular matrix (A3 domain of von Willebrand factor (vWF-A3) and fibrillar collagen) [Nishida et al., 2003]. It has recently been used to map interfaces of a 164 kDa protein-protein complex (B domain of protein A and intact IgG) [Nakanishi et al., 2002]. TCS method is very

robust and could tolerate wide sample conditions, which is exemplified by its use to map the binding faces of ligand proteins to membrane embedded in liposomes that were attached to solid beads (bead-linked proteoliposomes, composed of a potassium channel, KcsA, embedded in the lipid bilayer (KcsA-proteoliposomes), and used them to map the interaction interface between KcsA and its pore-blocking peptide, agitoxin2 (AgTx2)) [Yokogawa et al., 2005].

5.2.5.4 Mapping with dynamics

^{15}N relaxation experiments in NMR is used to describe dynamics in proteins. Use of these experiments to map the interface has a mixed message, since, in some cases the dynamics gets quenched and some others not. The former case is exemplified by the rigidification of the residues at the interface of the NSyp SH2 domain (tyrosine phosphatase) with peptide [Kay et al., 1998], shows increase in the order parameter for the bound state in comparison to its free state. Another example is, ^{15}N transverse relaxation rates for the free lipolyl domain and its complex with E1 component of the pyruvate dehydrogenase shows a loss of fast motion in the complex in comparison to the free form. However, the interaction of hydrophobic peptides with hydrophobic contact area of the PLCC SH2 domain (C-terminal SH2 domain of phospholipase C-γ1) shows no change in dynamics [Kay et al., 1996].

5.2.5.5 Mapping with amide-proton exchange (H-D Exchange)

Using amide-proton exchange one could map the interfaces of the protein. The solvent accessibility of the transiently unfolded areas allow the exchange of protons or deuterons, to take place on a timescales ranging from minutes to months. The formation of protein-protein complex is expected to protect the interface from such exchange. This method has been used for very high affinity interactions like antigen-antibody viz., epitopes of Der p 2 could be determined in the presence and absence of antibodies [Mueller et al., 2001].

5.2.5.6 Mapping with paramagnetics

Paramangnetic tools also appears to be a promising NMR method to identify protein-protein interface. This is based on the principle that if a protein is kept in a paramagnetic solution, only its surface residues will be affected by the paramagnetic line broadening. Interaction site between the catalytic domain of matrix metallo-proteinase 3 and its protein inhibitor TIMP was obtained by comparing the line

broadening of the amide-proton of mettaloproteinase by paramagnetic gadolinium-EDTA in the presence and absence of TIMP [Arumugam et al., 1998].

5.2.5.7 Mapping with site-specific spin labeling

TEMPO (2,2,6,6-tetramethylpiperidine-1-oxyl), a free stable radical nitroxide can interact specifically with cysteine groups on the protein surface [Jahnke et al., 2001]. The NMR line broadening of the resonances are proportional to the 1/r6 and hence useful distance restraints can be obtained from this data, which could be used for calculation of a structure. TE33-peptide complex model was calculated using this method [Scherf et al., 1995].

5.2.5.8 Relative positions of proteins using Residual Dipolar Couplings (RDC)

Dipolar coupling between the nuclei in a biomolecule averages to zero when it tumbles in a isotropic solution. But now suppose that the protein (biomolecule) is not dissolved in isotropic solution but rather in a media that leads to fractional alignment, then the dipolar coupling does not average to zero. The result is that small dipolar splittings (residual dipolar couplings, RDCs) are retained in the solution NMR spectra. The weak alignment can be obtained using magnetic fields, or in dilute solutions of bicelles or filamentous phages [Hansen et al., 1998; Tjandra and Bax, 1997; Tolman et al., 1997]. RDCs thus obtained has been used to obtain relative domain orientation [Fischer et al., 1999; Mueller et al., 2000] and or relative protein orientations in a multiprotein complex [Clore, 2000].

5.2.6 NMR based methods to calculate three-dimensional structures of protein complexes

The NMR structure determination is still mainly based on two pieces of information extracted from the NMR data, the determination of interatomic distances (NOEs), and dihedral angles. This information usually has to be completed by independent information from other sources describing the physical model including the covalent structure of the protein (Table 5.4). In addition, a number of other NMR-derived restraints can be measured that help to obtain a more refined structure.

Table 5.4: Restraints used for the NMR structure calculation

Data	Information
External restraints describing the physical model	
Covalent structure of amino acids and ligands (bond lengths, bond angles, dihedral angles, van der Waals radii, connectivity of atoms)	Possible positions of atoms in the conformational space
Amino acid sequence and possible ligands	Definition of the whole molecular complex
Disulfide bridges and other side-chain connections	Further restriction of the conformational space
Interatomic potentials (e.g., electrostatic potentials, Lennard–Jones potentials)	Further restriction of the conformational space
Homology models	Faster optimization
NMR derived data	
NOE	Interatomic distances
J-coupling	Dihedral angles
Chemical shift	Dihedral angles, interatomic distances, relative orientation of groups
Residual dipolar couplings	Interatomic distances and orientation
HNCO-type polarization transfer	Pairwise hydrogen bonding
Temperature dependence of amide chemical shifts	Hydrogen bonding
Hydrogen exchange rates	Hydrogen bonding
Non-averaged chemical shift anisotropy	Orientations of bond vectors, distance information

For large protein complexes, it is often very difficult to obtain enough NMR data to calculate high quality structures. It might become even more mandatory to obtain a structure with limited set of experimental data specially in situations like proteomics and high throughput screening. Hence a combination of computational and NMR data is now being successfully used for obtaining structures of large biomolecular complexes. Two broad ways of calculating structures are,

1. **Calculation of structures with limited set of experimental data:** Before proceeding to the docking procedure of protein-protein or protein-ligand, one has to know the 3D structure of the individual components. One such method is ROSETTA [Bowers et al., 2000] which uses NMR chemical shift data along with NOE data. Other methods involve use of NMR data with standard structure calculation procedures such as CNS or X-PLOR.

2. **Docking of proteins on the basis of experimental data:** NMR is particularly powerful in mapping the interfaces, allowing the study of weak and transient complexes that can be very difficult to be studied by other techniques. If the structures of the two individual components in a protein-protein interaction are known, then one could dock these two individual structures based on the experimentally obtained data. McCoy and Wyss [2002] first showed the use of CSP as a structural restraint for solving structures of protein complexes. Some

of the docking programs like BiGGER use experimental data only for scoring the final structures obtained. HADDOCK (High Ambiguity Driven biomolecular DOCKing) is the most widely used docking program, which uses ambiguous interaction restraints [Dominguez et al., 2003]. In the HADDOCK approach, CSP data are translated into ambiguous restraints to drive the docking process. The CSP interaction restraints can also be combined with RDC data allowing a better definition of the relative orientation of the components. In this thesis we have used CSP, CST, NOE and RDC data to obtain the structure of human Cdc37$_M$-Hsp90$_N$, protein-protein complex (Figure 5.7).

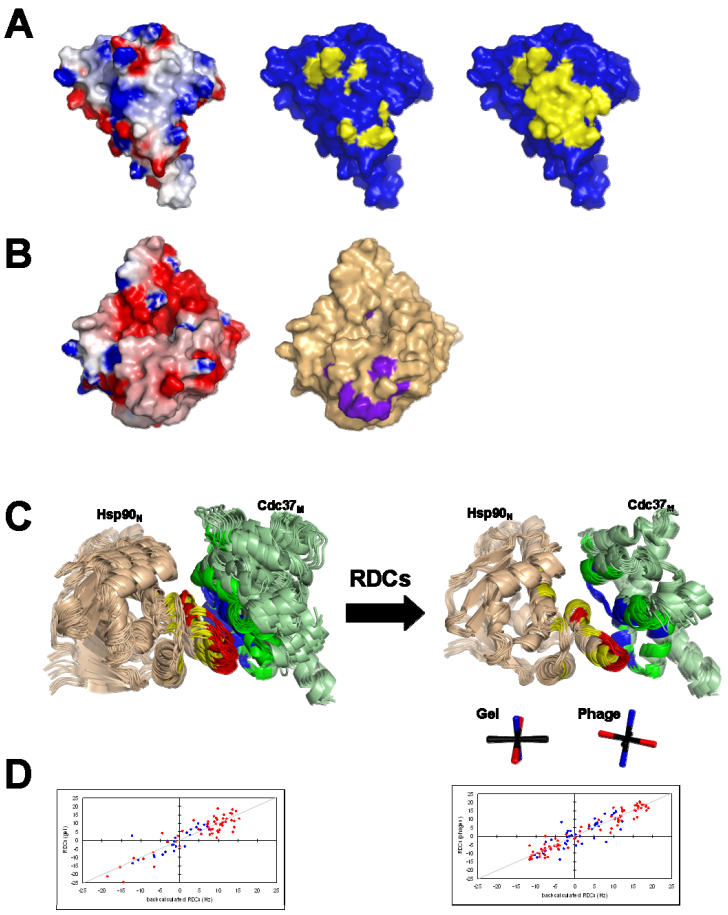


Figure 5.7: NMR data driven docking model of human Cdc37$_M$-Hsp90$_N$ complex. (A) Surface representation of Cdc37$_M$. (Left) The molecular surface is color-coded by electrostatic potential, as calculated with DELPHI. Potentials less than -10 kT are red, those greater than +10 kT are blue, and neutral potentials (0 kT) are white. The color in between is produced by linear extrapolation. (Middle) CSPs from Figure 5.3 ($>$ 0.15 ppm) in yellow mapped on to the structure. (Right) cross-saturation transfer experiment in which residues with intensity ratios less than 0.78 are mapped on to the surface (yellow). (B) Surface representation of human Hsp90$_N$ (1YES.pdb). (Left) The molecular surface is color-coded by electrostatic potential, as calculated with DELPHI. (Right) CSPs from Figure 5.3 ($>$ 0.1 ppm) in purple mapped on to the structure. (C) Bundle of the Cdc37$_M$-Hsp90$_N$ complex as obtained using HADDOCK, with and without using RDCs. For the calculation without RDCs (left), the ten best docking results for each of the two major HADDOCK clusters are shown and for the calculation with RDCs (right) the ten best docking results are shown. The active/passive restraints were used as defined in the materials and methods section and are colored respectively in red and yellow for Hsp90$_N$ or blue and dark-green for Cdc37$_M$. The RDC tensors for alignment in either polyacrylamide (PAA) gel or pf1 Phages are indicated below. (D) Correlation of the experimental versus back calculated RDCs for the free form X-ray structures of Cdc37$_M$ (blue) and Hsp90$_N$ (red). Protons were added to the X-ray structures using CNS 1.1 [Brünger, 1998] without using the RDCs. RDCs were back calculated using the program PALES [Zweckstetter and Bax, 2000] with the bestFit option. (Left) Correlation plot in PAA gel alignment medium with Cdc37$_M$ in blue (25 RDCs, Correlation co-efficient of 0.822) and Hsp90$_N$ in red (52 RDCs, Correlation co-efficient of 0.903). (Right) Correlation plot in pf1 phages alignment medium with Cdc37 in blue (42 RDCs, Correlation co-efficient of 0.816) and Hsp90$_N$ in red (79 RDCs, Correlation co-efficient of 0.964).

6 Mutational and Folding Studies of Protein kinase A by NMR spectroscopy

6.1 Research Article

Folding and Activity of cAMP-Dependent Protein kinase Mutants

Thomas Langer, Sridhar Sreeramulu, Martin Vogtherr, Bettina Elshorst,
Marco Betz, Ulrich Schieborr, Krishna Saxena, Harald Schwalbe,
FEBS Letters, 2005, 579, 4049-4054.

In this article we used NMR to study the folding and activity of cAMP-dependent protein kinase (PKA, a surrogate of oncogenic kinase AKT/PKB) mutants.

AKT/PKB is a oncogenic protein kinase, which is a very well validated drug target. Therefore, obtaining specific small molecule inhibitors for Akt/PKB is highly desirable. A major obstacle on this way is the production of large quantities of pure and active Akt/PKB protein for structural studies, since it cannot be expressed in *Escherichia coli* in a functional form. Furthermore, Akt/PKB has to be phosphorylated on the activation loop as well as in a regulatory sequence, called the hydrophobic motif . Since both PKA and Akt/PKB belong to the AGC-family of protein kinases and share high sequence homology, PKA has been used as a surrogate kinase for Akt/PKB.

PKA has four phosphorylation sites and undergoes autophosphorylation when expressed in *E. coli*. But the autophosphorylation is not uniform and hence results in mixture of differently phosphorylated isoforms, leading to requirement of an additional purification step and also making it not very suitable for precision-oriented structural studies. In this contribution we made several phospho-mutants of PKA, accessed their folding by using TROSY NMR and also compared their activity with wild-type.

The author of this thesis has performed all molecular biology work and NMR analysis, in part together with Dr. Thomas Langer and Dr. Martin Vogtherr.

Folding and activity of cAMP-dependent protein kinase mutants

Thomas Langer*, Sridhar Sreeramulu, Martin Vogtherr, Bettina Elshorst, Marco Betz, Ulrich Schieborr, Krishna Saxena, Harald Schwalbe*

Johann Wolfgang Goethe-University Frankfurt, Institute for Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance, Marie Curie Strasse 11, D-60439 Frankfurt am Main, Germany

Received 9 May 2005; revised 7 June 2005; accepted 10 June 2005

Available online 6 July 2005

Edited by Miguel De la Rosa

Abstract The catalytic subunit of cAMP-dependent protein kinase (PKA) can easily be expressed in *Escherichia coli* and is catalytically active. Four phosphorylation sites are known in PKA (S10, S139, T197 and S338), and the isolated recombinant protein is a mixture of different phosphorylated forms. Obtaining uniformly phosphorylated protein requires separation of the protein preparation leading to significant loss in protein yield. It is found that the mutant S10A/S139D/S338D has similar properties as the wild-type protein, whereas additional replacement of T197 with either E or D reduces protein expression yield as well as folding propensity of the protein. Due to its high sequence homology to Akt/PKB, which cannot easily be expressed in *E. coli*, PKA has been used as a surrogate kinase for drug design. Several mutations within the ATP binding site have been described to make PKA even more similar to Akt/PKB. Two proteins with Akt/PKB-like mutations in the ATP binding site were made (PKAB6 and PKAB8), and in addition S10, S139 and S338 phosphorylation sites have been removed. These proteins can be expressed in high yields but have reduced activity compared to the wild-type. Proper folding of all proteins was analyzed by 2D ^{1}H, ^{15}N-TROSY NMR experiments.

Keywords: Nuclear magnetic resonance; Protein kinase A; PKA; Akt; Protein phosphorylation

1. Introduction

Protein kinases are a major protein class that regulate a multitude of cellular processes. This is achieved by catalyzing the transfer of the γ-phosphate from ATP to the target protein, thereby changing its activity. Due to their indispensable function in cellular signaling, deregulation of certain protein kinases result in disease. Inhibition of these deregulated protein kinases by small molecule inhibitors is pharmaceutically relevant [1,2]. Successful examples of this strategy are Gleevec (imatinib), which inhibits cAbl [3], or Iressa (gefitinib), which targets the kinase domain of EGFR [4]. Akt/PKB is a protein kinase that has been found frequently in a constitutively active form in human cancers. The central role of Akt/PKB is well validated [5,6]. Therefore, obtaining specific small molecule inhibitors for Akt/PKB is highly desirable. A major obstacle on this way is the production of large quantities of pure and active Akt/PKB protein for structural studies, since it cannot be expressed in *Escherichia coli* in a functional form. Furthermore, Akt/PKB has to be phosphorylated on the activation loop as well as in a regulatory sequence, called the hydrophobic motif [6]. Since both PKA and Akt/PKB belong to the AGC-family of protein kinases and share high sequence homology, PKA has been used as a surrogate kinase for Akt/PKB. To make PKA more similar to Akt/PKB distinct mutations within the ATP-binding region have successfully been produced [7,8]. PKA has also been used as a surrogate for Rho-kinase [9].

Nearly all protein kinases are regulated by phosphorylation themselves. The most critical part hereby lies in the so-called activation loop. Upon phosphorylation of a certain residue within the activation loop, the protein kinase active site is reoriented to allow proper ATP and substrate binding. When expressed in *E. coli*, PKA is catalytically active and undergoes autophosphorylation. There are four phosphorylation sites within PKA: S10, S139, T197 and S338. The recombinantly expressed protein is a mixture of differently phosphorylated isoforms with 2, 3 or 4 phosphate groups; the protein lacks a unique phosphorylation pattern [10–12]. Phosphorylation of T197 in the activation loop is essential for activity. Changing T197 to alanine results in a folded, but unstable and inactive protein. Phosphorylation of S338 has no influence on activity but enhances protein stability [11,13]. For PKA, phosphorylation of residue S10 has not been found in preparations from mammalian sources and a physiological role of S10 phosphorylation has only recently been proposed [14]. Also, PKA from mammalian tissues with phosphorylated S139 has not been observed and there are currently no clues for any functional relevance. The extent of phosphorylation of S10 and S139 in *E. coli* is obviously dependent on the expression conditions.

To avoid the need for separating differently phosphorylated PKA isoforms we wondered whether it is possible to produce active protein where the catalytic non-relevant phosphorylation sites have been removed or even a "dephospho"-PKA lacking all phosphorylation sites. Since phosphorylation of S10 disorders the N-terminal helix [14] we replaced S10 with alanine, whereas S139 and S338 were changed to aspartate (PKA-3P). T197 was either replaced by aspartate or glutamate resulting in PKA-4P(D) and PKA-4P(E).

Since PKA serves as a Akt/PKB surrogate kinase in the drug discovery process, two PKA constructs with 6 and 8

*Corresponding authors. Fax: +49 69 798 29515.
E-mail addresses: t.langer@nmr.uni-frankfurt.de (T. Langer), schwalbe@nmr.uni-frankfurt.de (H. Schwalbe).

0014-5793/$30.00
doi:10.1016/j.febslet.2005.06.015

mutations in the ATP binding region were made (PKAB6: Q84E/V104T/V123A/L173M/Q181K/F187L and additionally S53T/L82K for PKAB8). Finally, the S10A, S139D and S338D mutations were also introduced in the PKAB constructs yielding PKAB6-3P and PKAB8-3P, respectively. The folding of all expressed proteins was checked by 2D ^{1}H, ^{15}N TROSY nuclear magnetic resonance spectroscopy (NMR) and the activity of the proteins were assessed by standard assays for enzymatic activity.

2. Materials and methods

2.1. Cloning, expression and purification of PKA

PKA was expressed in *E. coli* strain Bl21 (DE3)/pLysS (Novagen) as described [13]. The DNA coding for PKA was originally cloned from CHO cells (references cited in [13]). Site specific mutagenesis was done following the Quick Change protocol (Stratagene). All constructs have been verified by sequencing. To obtain uniformly ^{15}N labeled protein, cells were grown in M9 minimal-medium containing $^{15}NH_4Cl$ (Eurisotop, Saarbrücken, Germany) as the sole nitrogen source. For all constructs, cell growth and protein expression was performed under the same conditions. Cells were grown in the presence of appropriate antibiotics at 37 °C until the OD_{600} reached 0.6 and then recombinant protein production was initiated by adding 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside). Induction was carried out for 4 h at 28 °C and cells were harvested by centrifugation for 15 min at 4500*g*. PKA proteins were purified with the aid of a His-tag as described in detail previously [13]. The His-tag was removed using a Tobacco Etch virus (Tev)-cleavage site.

2.2. NMR spectroscopy

All protein samples were measured in the same buffer (150 mM NaCl, 5 mM DTT, 50 mM HEPES, pH 6.5, 1 mM TSP, in 90% H_2O/10% D_2O). Protein concentration was 0.25 mM for all preparations except for the "dephospho"-forms PKA-4P(D) and PKA-4P(E), where the protein was concentrated to 0.1 mM. 800 MHz 1D-Watergate and 2D-TROSY [15] experiments (128 T_1 increments) were acquired at 298 K on a Bruker DRX800 spectrometer.

2.3. Enzymatic assay

Enzymatic activity of PKA proteins was checked with the aid of a luminescent kinase assay following the manufacturer's recommendations (Kinase-Glo, Promega, Madison, WI, USA). Kemptide (Promega) was used as substrate. The assay were performed in 96 well plates with 50 mM HEPES, pH 7.5, 50 mM NaCl, 2 mM DTT, 10 mM $MgCl_2$, 2.5 μM ATP and 5 μM substrate peptide. The amount of employed PKA proteins was 74 nM in all assays. Samples were incubated for 90 min at room temperature before addition of the Kinase-Glo solution. After 20 min, luminescence was recorded with a Veritas microplate luminometer (Turner BioSystems, Sunnyvale, CA, USA). Data were analyzed using GraphPad Prism (GraphPad Software, Inc., San Diego, CA, USA).

3. Results

All proteins were expressed under identical conditions at 28 °C. For the "dephospho"-PKA-4P(D) and PKA-4P(E) proteins, most of the protein was expressed as insoluble inclusion bodies. However, the yield of soluble protein was 1.5 and 1.0 mg/l M9 minimal medium for PKA-4P(D) and PKA-4P(E), respectively. Typical expression yields for wild-type PKA are 10 mg/l. For PKAB6 and PKAB8 similar amounts of 10 mg/l were obtained. Interestingly, even higher yields were obtained from the corresponding −3P forms (15–20 mg/l). Except for the "dephospho"-PKA-4P proteins, all PKA mutants could easily be concentrated up to 0.5 mM without precipitation. Therefore, PKA-4P(D) and PKA-4P(E) were only concentrated to 0.1 mM for NMR measurements. The phosphorylation sites of PKA are depicted in Fig. 1.

In order to evaluate whether the expressed proteins are correctly folded, 2D ^{1}H, ^{15}N TROSY NMR spectra were recorded and compared with the corresponding spectrum of wild-type PKA, for which NMR backbone assignment of PKA has recently been published by us [13]. As shown in Fig. 2(a), PKA-3P is properly folded. Some peaks are shifted, which is obviously a result of the mutations. Concerning PKA-4P(D) and PKA-4P(E) samples, there are certain amounts of unfolded protein or only partially folded protein (Fig. 2(b) and (c)).

Several PKAB chimeras have been characterized both functionally and structurally [7,8]. We have extended this principle by introducing further mutations leading to the new chimeras PKAB6 and PKAB8 (Fig. 3). S53T and L82K are outside of the ATP-binding pocket and may allow for designing small molecule inhibitors that extend beyond the ATP-binding pocket and therefore confer more selectivity. Since PKA-3P can easily be expressed and is stable, the corresponding mutations were introduced in PKAB6 and PKAB8, resulting in PKAB6-3P and PKAB8-3P, respectively. Again, correct folding of all PKAB proteins was confirmed by 2D ^{1}H, ^{15}N-NMR spectroscopy. The different PKA mutants display shifted peaks but the overall peak pattern is similar compared to wild-type PKA (Fig. 4).

Catalytic activity was monitored with a luminescent kinase assay. The activity of wild-type PKA was set to 100%. The results are shown in Fig. 5. PKA-3P is as active as wild-type PKA and even the "dephospho"-PKA-4P(D) and PKA-4P(E) proteins display 57% and 39%, activity, respectively, compared to wild-type PKA. PKAB6 and PKAB8 have reduced activity (15% and 10%, respectively). These results were unexpected, but as a consequence of the introduced

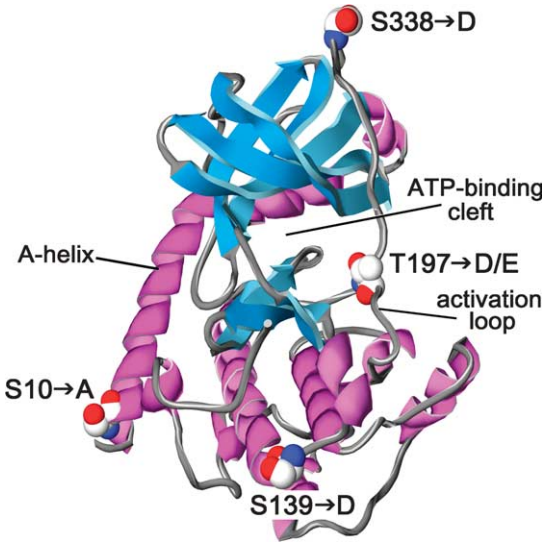


Fig. 1. Structure of the PKA catalytic subunit. The structure is drawn from pdb-file 1CTP with Swiss pdb-viewer. The side chains of the amino acids that can be phosphorylated are shown as spheres. A unique feature of PKA is the A-helix. The ATP-binding cleft lies between the N- and C-terminal lobe.

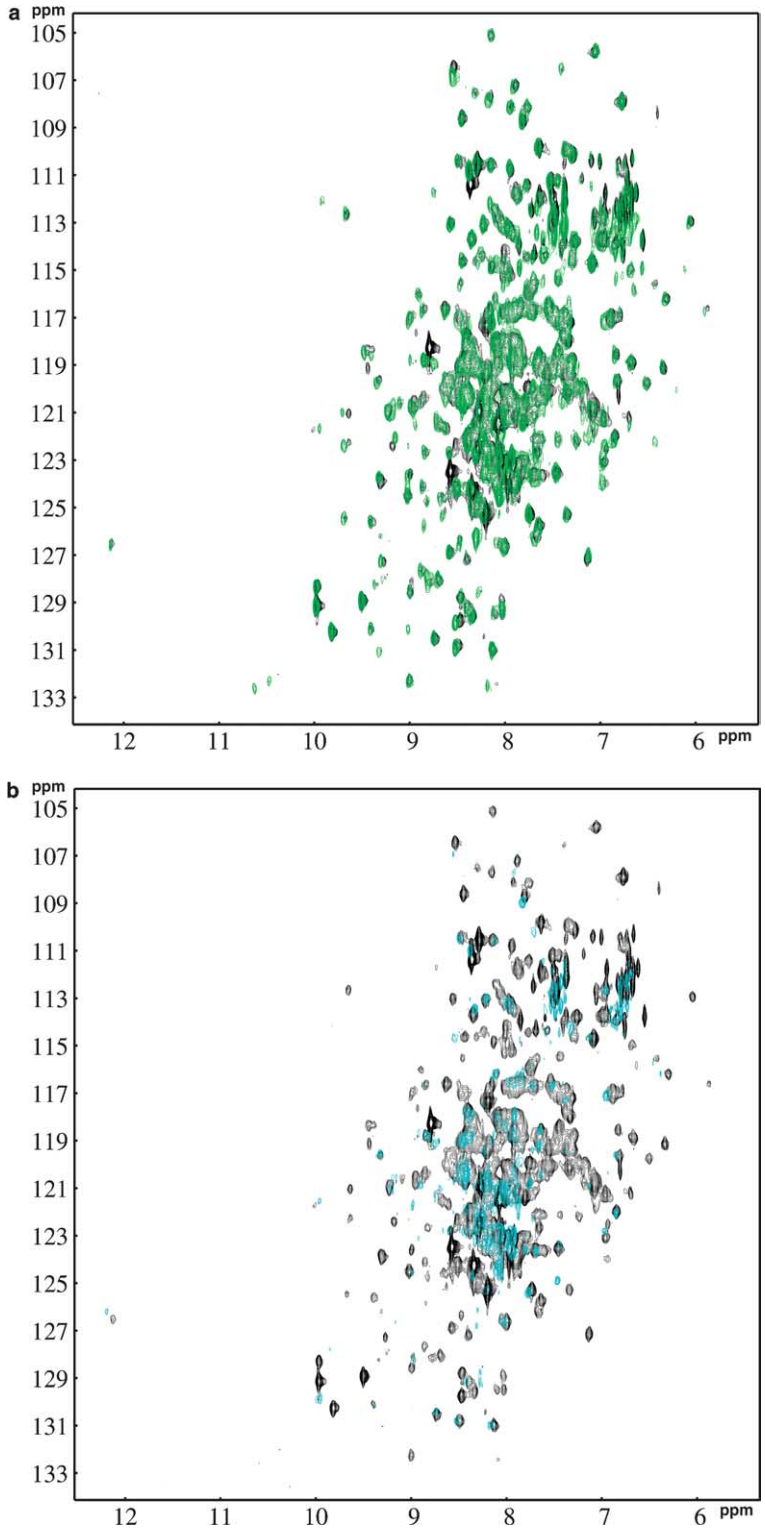


Fig. 2. Comparison of TROSY spectra of wild-type PKA (black) with "dephospho"-PKA mutants. (a) PKA-3P (green), (b) PKA-4P(D) (blue) and (c) PKA-4P(E) (red). See text for details.

mutations, substrate binding might be compromised. Interestingly, PKAB6-3P and PKAB8-3P were more active as the corresponding phosphorylateable forms (28% and 43% compared to wild-type PKA for PKAB6-3 and PKAB8-3, respectively).

4. Discussion

PKA containing S10A, S139D and S338D mutations can easily be expressed in *E. coli*, the protein is folded and as active as wild-type PKA. Replacing the additional phosphorylation site

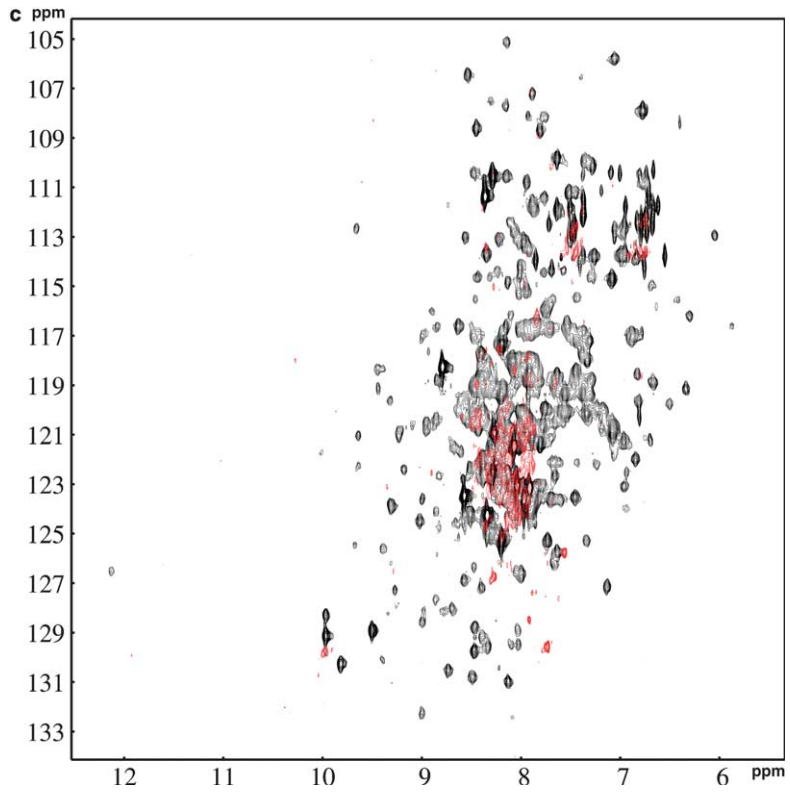


Fig. 2 (*continued*)

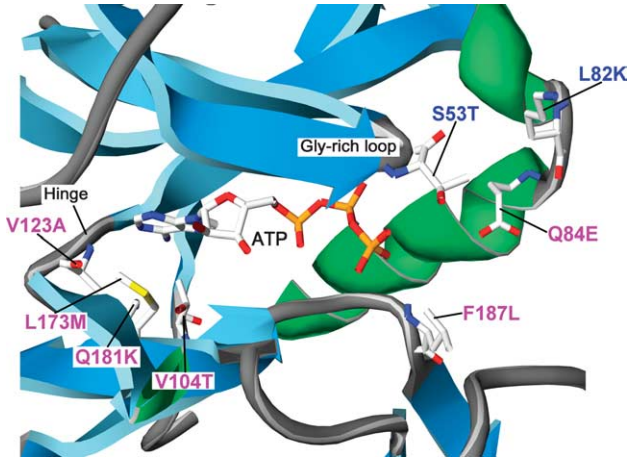


Fig. 3. ATP binding site of PKA with bound ATP. The structure is drawn from pdb-file 1ATP with Swiss pdb-viewer. The hinge region, connecting the N- and C-terminal lobe of the protein, is indicated. Mutations to enhance Akt/PKB similarity are displayed. Mutations depicted in magenta belong to PKAB6 and the additional mutations (indicated in blue) S53T and L82K pertain to PKAB8.

T197 by either D or E yields poorly soluble protein. Similar results have been described for the single mutants PKA T197D and PKA T197E [11]. As reported previously PKA T197E was found to be nearly inactive whereas PKA T197D has an activity comparable to the wild-type [11]. By contrast, we have measured significant, albeit lower activity for both the PKA-4P(D) and the PKA-4P(E) mutants (57% and 39% activity, respectively) relative to the wild type. The physiological role of S10 has been elucidated recently [14]. The N-terminus of PKA folds into the so-called A-helix which is unique for PKA. Mammalian PKA is myristoylated on Gly1, and Asn2 is deamidated. These modifications are absent if PKA is expressed in *E. coli*. Phosphorylation of S10 is presumed to alter the structure and hence regulate the localization of PKA. It does so by inducing destabilization and disorder of the first nine amino acids in S10-phosphorylated PKA, which otherwise adopts an α-helical conformation in S10 unphosphorylated PKA [14]. In an earlier report, Yonemoto et al. [11] describe that PKA with only the S10A mutation yields poor soluble protein. However, the PKA-3P mutant with additional S139D and S338D mutations is soluble, properly folded and can be expressed in high yields. Insertion of either aspartate or glutamate instead of T197, which lies within the activation loop, resulted in poorly soluble protein of lower quality. These observations might be explained in that way, that S139D and S338D mutations have a positive effect on protein folding. This would be in accordance with the protein stabilising effect of S338 phosphorylation. Replacing S338 by alanine yields a more unstable, but fully active protein [13]. In contrast, changing S139 to alanine does not display adverse effects [11].

4.1. PKAB chimeras

PKA has been used as a surrogate kinase for Akt/PKB and Rho [7–9]. To enhance the usefulness of PKA as Akt/PKB surrogate, several mutations have been introduced within the ATP binding site that correspond to Akt/PKB residues. Here, we extend such PKAB chimeras with the introduction of S53T, L82K and V104T mutations. S53 and L82 that are outside the ATP-binding-pocket may be helpful in the design of inhibitors

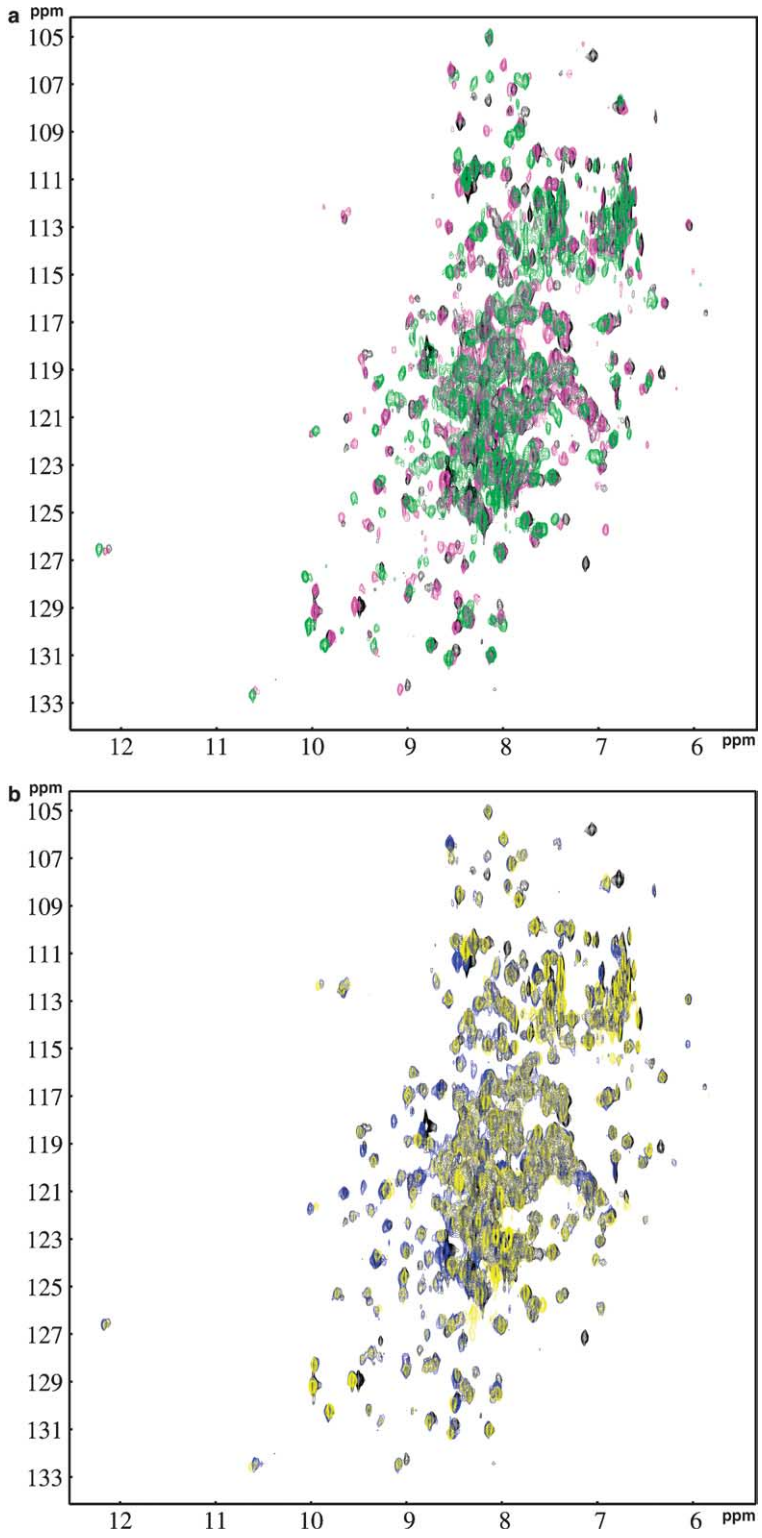


Fig. 4. Comparison of TROSY spectra of wild-type PKA (black) with the PKAB chimerical proteins. (a) PKAB6 (magenta), PKAB6-3P(green) and (b) PKAB8 (blue), PKAB8-3P (yellow).

that protrude from the ATP-binding pocket. PKAB6 and PKAB8 can easily be expressed in *E. coli* and are properly folded. Both mutants are less active than wild-type PKA. This may be due to reduced affinity for ATP or the substrate peptide as a consequence of the introduced mutations. A 12-fold reduced affinity for ATP has been reported for the PKA Q84E/V123A/L173M/F187L mutant. The increase in yield and activity for mutations at the phosphorylation sites (S10A, S139D and S338D) is remarkable. This allows expression of PKA and PKAB chimeras with only the T197 phosphorylation site remained. Since no further phosphorylation sites are present, no mixtures of differently phosphorylated

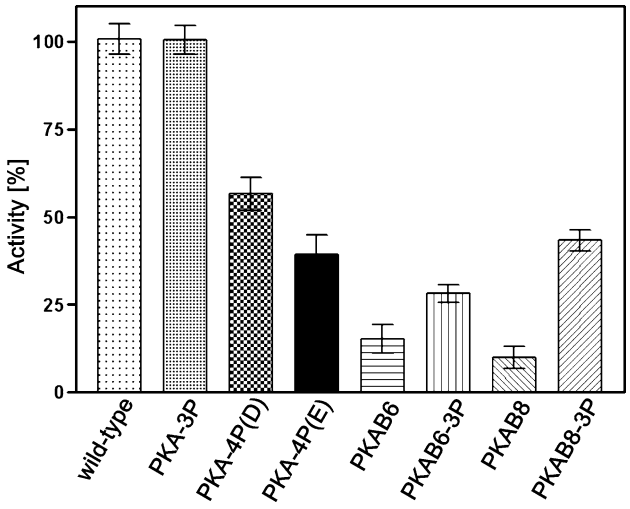


Fig. 5. Enzymatic activity of wild type PKA compared with "dephospho"-PKA, and PKAB chimerical proteins. Error bars represent the standard error of the mean (SEM). The activity of wild-type PKA is set as 100% and activity of PKA mutants are related to that. For experimental details see Section 2.

isoforms occur and consequently there is no need for further separation. This should be highly advantageous when the protein is used for X-ray crystallography and functional studies.

References

[1] Fabbro, D., Ruetz, S., Buchdunger, E., Cowan-Jacob, S., Fendrich, G., Liebetanz, J., Mestan, J., O'Reilly, T., Traxler, P., Chaudhuri, B., Fretz, H., Zimmermann, J., Meyer, T., Caravatti, G., Furet, P. and Manley, P. (2002) Protein kinases as targets for anticancer targets: from inhibitors to useful drugs. Pharmacol. Therap. 93, 79–98.

[2] Noble, M.E.M., Endicott, J.A. and Johnson, L.N. (2004) Protein kinase inhibitors: insight into drug design from structure. Science 303, 1800–1805.

[3] Ren, R. (2005) Mechanisms of BCR-ABL in the pathogenesis of chronic myelogenous leukaemia. Nat. Rev. Cancer 5, 172–183.

[4] Birnbaum, A. and Ready, N. (2005) Gefitinib therapy for non-small cell lung cancer. Curr. Treat. Options Oncol. 6, 75–81.

[5] Brognard, J., Clark, A.S., Ni, Y. and Dennis, P.A. (2001) Akt/protein kinase B is constitutively active in non-small cell lung cancer cells and promotes cellular survival and resistance to chemotherapy and radiation. Cancer Res. 61, 3986–3997.

[6] Hanada, M., Feng, J. and Hemmings, B.A. (2004) Structure, regulation and function of PKB/AKT – a major therapeutic target. Biochim. Biophys. Acta 1697, 3–16.

[7] Gassel, M., Breitenlechner, C.B., Ruger, P., Jucknischke, U., Schneider, T., Huber, R., Bossemeyer, D. and Engh, R.A. (2003) Mutants of protein kinase A that mimic the ATP-binding site of protein kinase B (AKT). J. Mol. Biol. 329, 1021–1034.

[8] Breitenlechner, C.B., Friebe, W.G., Brunet, E., Werner, G., Graul, K., Thomas, U., Kunkele, K.P., Schafer, W., Gassel, M., Bossemeyer, D., Huber, R., Engh, R.A. and Masjost, B. (2005) Design and crystal structures of protein kinase B-selective inhibitors in complex with protein kinase A and mutants. J. Med. Chem. 48, 163–170.

[9] Breitenlechner, C., Gassel, M., Hidaka, H., Kinzel, V., Huber, R., Engh, R.A. and Bossemeyer, D. (2003) Protein kinase A in complex with Rho-kinase inhibitors Y-27632, Fasudil, and H-1152P: structural basis of selectivity. Structure 11, 1595–1607.

[10] Herberg, F.W., Bell, S.M. and Taylor, S.S. (1993) Expression of the catalytic subunit of cAMP-dependent protein kinase in *Escherichia coli*: multiple isozymes reflect different phosphorylation states. Prot. Eng. 6, 771–777.

[11] Yonemoto, W., McGlone, M.L., Grant, B. and Taylor, S.S. (1997) Autophosphorylation of the catalytic subunit of cAMP-dependent protein kinase in *Escherichia coli*. Prot. Eng. 10, 915–925.

[12] Seifert, M.H., Breitenlechner, C.B., Bossemeyer, D., Huber, R., Holak, T.A. and Engh, R.A. (2002) Phosphorylation and flexibility of cyclic-AMP-dependent protein kinase (PKA) using ^{31}P NMR spectroscopy. Biochemistry 41, 5968–5977.

[13] Langer, T., Vogtherr, M., Elshorst, B., Betz, M., Schieborr, U., Saxena, K. and Schwalbe, H. (2004) NMR backbone assignment of a protein kinase catalytic domain by a combination of several approaches: application to the catalytic subunit of cAMP-dependent protein kinase. Chembiochem 5, 1508–1516.

[14] Breitenlechner, C.B., Engh, R.A., Huber, R., Kinzel, V., Bossemeyer, D. and Gassel, M. (2004) The typically disordered N-terminus of PKA can fold as a helix and project the myristoylation site into solution. Biochemistry 43, 7743–7749.

[15] Pervushin, K., Riek, R., Wider, G. and Wüthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc. Natl. Acad. Sci. USA 94, 12366–12371.

7 Resonance Assignment and Structural Studies of the Human Cdc37-Hsp90 Complex by NMR spectroscopy

7.1 Research Article

^{1}H, ^{13}C and ^{15}N Backbone Resonance Assignment of the Hsp90 Binding Domain of Human Cdc37

Sridhar Sreeramulu, Jitendra Kumar, Christian Richter, Martin Vogtherr, Krishna Saxena, Thomas Langer, Harald Schwalbe,
Journal of Biomolecular NMR, 2005, 32, 262.

7.2 Research Article

The Human Cdc37-Hsp90 Complex Studied by Heteronuclear NMR Spectroscopy

Sridhar Sreeramulu, Hendrik R. A. Jonker, Thomas Langer, Christian Richter, C. Roy D. Lancaster, Harald Schwalbe,
Journal of Biological Chemistry, 2009, 284, 3885-3896.

The two articles present liquid-state NMR studies of the 40 kDa human Cdc37-Hsp90, protein-protein complex. It also sets a model example of how NMR can be used to study such high molecular weight complexes.

The first contribution describes the complete backbone resonance assignment of the human Hsp90 binding domain of Cdc37.

In the second contribution, the structure of the 40 kDa Cdc37-Hsp90 complex was determined by using a combination of NMR, X-ray and data driven computational docking. The X-ray structure of the human Hsp90 binding domain of Cdc37, was solved in its monomeric form. Further, it also describes how residual dipolar couplings (RDCs) was used to arrive at the more refined structure of the complex. It also emphasizes how NMR data was used to identify the 'Hot Spot' in the large interaction interface of the protein-protein complex, which was further confirmed by mutational studies.

The author of this thesis has performed all the molecular biology and NMR experiments and analysis. X-ray crystal structure of Cdc37 was solved together with Prof. Dr. C. Roy D. Lancaster and the structure of the complex together with Dr. Hendrik R. A. Jonker.

Journal of Biomolecular NMR (2005) 32: 262

Letters to the Editor

NMR assignment of the novel *Helicobacter pylori* protein JHP1348

DOI 10.1007/s10858-005-8529-0

Helicobacter pylori is a pathogenic bacterium that infects a large proportion of the population worldwide. It is the causative agent for most cases of gastritis, gastric ulcer, and can lead to gastric cancer. Although the genomes of two strains of *H. pylori* have been sequenced, approximately one third of the open reading frames have no known homologs and no assigned functions (Tomb et al., 1997). Using a structural genomics paradigm, we searched the genome of *H. pylori* strain J99 for open reading frames encoding for functionally uncharacterized proteins with properties that would make their structural determination feasible by solution NMR techniques. JHP1348 is a putative periplasmic protein of 113 amino acids. Nearly all backbone and sidechain ^{1}H, ^{13}C, and ^{15}N resonances have been assigned using 2D and 3D heteronuclear NMR experiments (Sattler et al., 1999). Only C' resonance assignments for L18, which precedes a proline, and E111, which is the C-terminal residue, are missing. Over 87% of sidechain ^{1}H, ^{13}C, and ^{15}N resonances have been assigned. The chemical shifts have been deposited in the BioMagResBank under accession number 6640.

References: Tomb et al. (1997) *Nature*, **388**, 539–547; Sattler et al. (1999) *Prog. NMR Spectrosc.*, **34**, 93–158.

Brendan N. Borin[a], Andrei Popescu[a], Andrzej M. Krezel[a,*]

[a]*Department of Biological Sciences and Center for Structural Biology, Vanderbilt University, Nashville, TN, 37232, U.S.A*

*To whom correspondence should be addressed. E-mail: Andrzej.M.Krezel@Vanderbilt.edu

Supplementary material to this paper is available in electronic format at http://dx.doi.org/10.1007/s10858-005-8529-0.

^{1}H, ^{13}C and ^{15}N backbone resonance assignment of the Hsp90 binding domain of human Cdc37

DOI 10.1007/s10858-005-8530-7

Cdc37 belongs to the cohort of cochaperones that assist Hsp90 to function correctly and is mainly associated with Hsp90. Most of the client proteins interacting with Hsp90/Cdc37 are crucial elements of signaling pathways inside the cell. A predominant group herein are protein kinases. Hence, Cdc37 was assigned as a kinase targeting subunit of Hsp90 (MacLean and Picard, 2003). Cdc37 can be dissected into three different domains. An N-terminal client binding domain, a middle domain binding Hsp90 and a C terminal domain without yet known function. Recently, a crystal structure of the middle domain bound to the N-terminal ATP-binding domain of Hsp90 became available (Roe et al. 2004). The Hsp90 binding domain of human Cdc37 comprising amino acids 147–276 were overexpressed in *E. coli* and 3D heteronuclear NMR experiments with ^{2}H, ^{13}C, ^{15}N and 2D NMR experiments with uniformly ^{15}N labelled protein were used for assignment. For the 130 residue fragment of Cdc37 the backbone assignments are essentially complete with exception of Pro 230. From the triple resonance experiments, further backbone and non-aromatic side chain assignments were made to the following extents: 98% of C_{α}; 86% of H_{α}; 97% of C_{β}, 73% of H_{β}; and 96% of CO. BMRB deposit with accession number BMRB-6586.

References: MacLean and Picard (2003) *Cell Stress Chaperones*, **8**, 114–119; Roe et al. (2004) *Cell*, **116**, 87–96.

Sridhar Sreeramulu[a], Jitendra Kumar[a], Christian Richter[a], Martin Vogtherr[a], Krishna Saxena[a], Thomas Langer[a], Harald Schwalbe[a,*]

[a]*Institut für Organische Chemie und Chemische Biologie, Zentrum für Biologische Magnetische Resonanz, Johann Wolfgang Goethe-Universität Frankfurt, Marie Curie Straße 11, D-60439, Frankfurt*

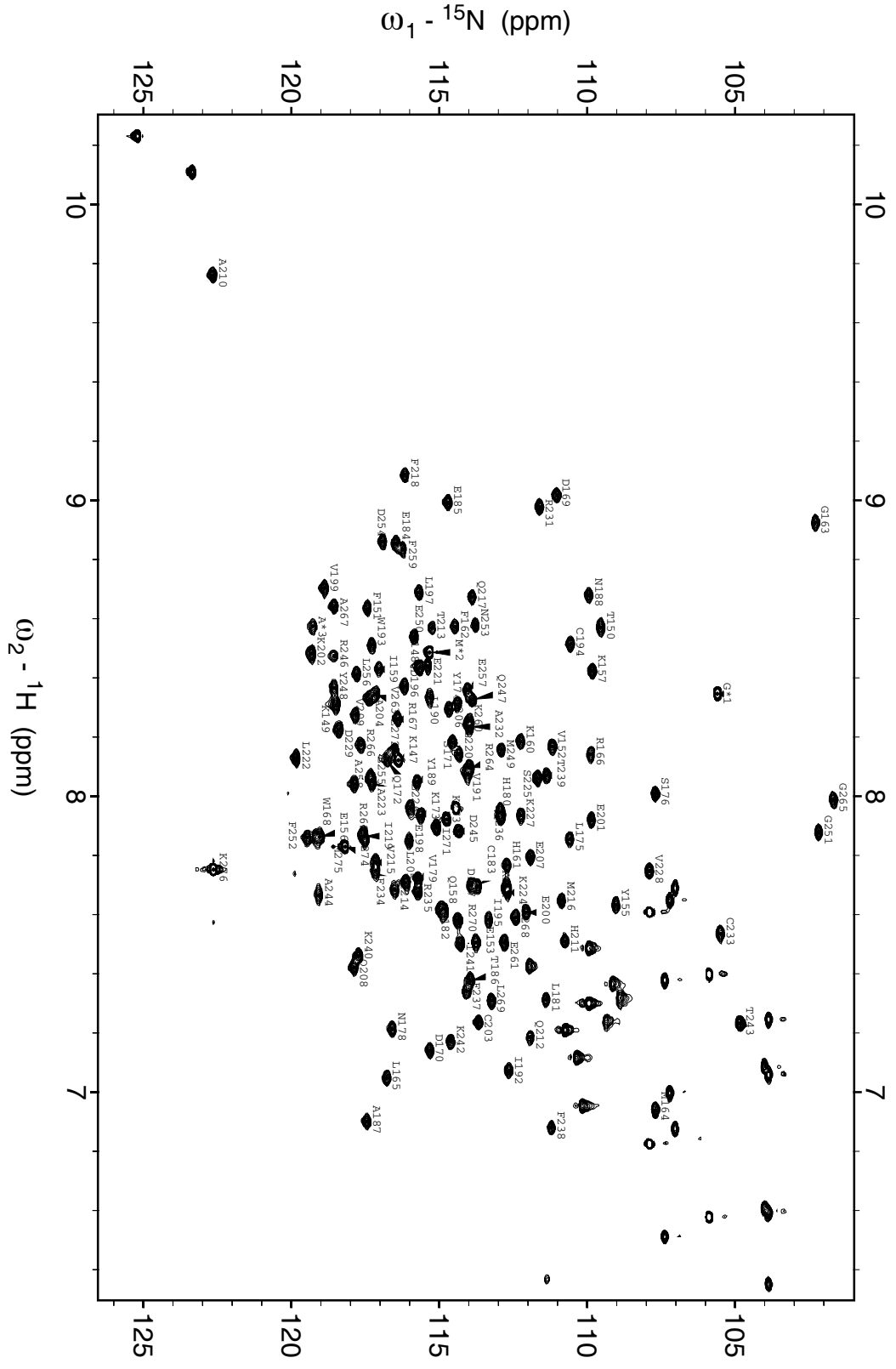
*To whom correspondence should be addressed. E-mail: schwalbe@nmr.uni-frankfurt.de

Supplementary material to this paper is available in electronic format at http://dx.dio.org/10.1007/s10858-005-8530-7

Figure Legend

Figure 1: ^{1}H-^{15}N spectrum of the Hsp90 binding domain of human Cdc37, 25ºC, pH 5.5

Supplementary Figure 1. Sreeramulu et al.



THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 6, pp. 3885–3896, February 6, 2009

The Human Cdc37·Hsp90 Complex Studied by Heteronuclear NMR Spectroscopy*

Received for publication, August 29, 2008, and in revised form, November 18, 2008 Published, JBC Papers in Press, December 10, 2008, DOI 10.1074/jbc.M806715200

Sridhar Sreeramulu[‡], Hendrik R. A. Jonker[‡], Thomas Langer[‡1], Christian Richter[‡], C. Roy D. Lancaster[§¶], and Harald Schwalbe[‡2]

From the [‡]Institute for Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance, Johann Wolfgang Goethe-University, Max-von-Laue-Strasse 7, Frankfurt am Main D-60438, the [§]Department of Molecular Membrane Biology, Max Planck Institute of Biophysics, Max-von-Laue-Strasse 3, Frankfurt am Main D-60438, and the [¶]Department of Structural Biology, Saarland University, Faculty of Medicine, Building 60, Homburg D-66421, Germany

The cell division cycle protein 37 (Cdc37) and the 90-kDa heat shock protein (Hsp90) are molecular chaperones, which are crucial elements in the protein signaling pathway. The largest class of client proteins for Cdc37 and Hsp90 are protein kinases. The catalytic domains of these kinases are stabilized by Cdc37, and their proper folding and functioning is dependent on Hsp90. Here, we present the x-ray crystal structure of the 16-kDa middle domain of human Cdc37 at 1.88 Å resolution and the structure of this domain in complex with the 23-kDa N-terminal domain of human Hsp90 based on heteronuclear solution state NMR data and docking. Our results demonstrate that the middle domain of Cdc37 exists as a monomer. NMR and mutagenesis experiments reveal Leu-205 in Cdc37 as a key residue enabling complex formation. These findings can be very useful in the development of small molecule inhibitors against cancer.

The 90-kDa heat shock protein (Hsp90)[3] is an abundant and essential protein, accounting for 1–2% of the total protein amount in unstressed eukaryotic cells. A 2- to 10-fold higher expression level of Hsp90 is found in cancer cells and virally transformed cells (1–3), suggesting a crucial role of Hsp90 for growth and/or survival of tumor cells. Hsp90 is a protein chaperone that helps a variety of client proteins to adopt their native conformation. Several co-chaperones such as Hop/Sti1, FKBP52, cyclophilin 40, and p50/Cdc37 are identified to be associated with Hsp90. These co-chaperones are required for binding to diverse client proteins and controlling the ATPase cycle of Hsp90 (4–6). Hsp90 is composed of three domains: an N-terminal domain of ~25-kDa, a middle domain of ~35-kDa, and a C-terminal domain of ~10-kDa (7). The ATP-binding site of Hsp90 is located in the N-terminal domain. Low molecular weight compounds that interfere with ATP-binding such as geldanamycin or radicicol inhibit the Hsp90 function by competing for the interaction site.

Many of the client proteins of Hsp90 are crucial elements of the cellular signal transduction pathway. Protein kinases represent the largest class of Hsp90 client proteins, and many have been shown to cause cancer when deregulated. These protein kinases depend on Hsp90 for their proper folding and functioning suggesting that Hsp90 itself is a valid anti-cancer target of pharmaceutical interest (8).

Hsp90 does not directly bind to protein kinases, but this interaction is mediated by the co-chaperone Cdc37/p50. Hence, Cdc37 has also been dubbed as a kinase targeting subunit of the Hsp90 machinery (9). Interestingly, Cdc37 is also able to act as a chaperone by itself independent of Hsp90 (10).

The cell division cycle protein 37 (Cdc37) was first identified in *Saccharomyces cerevisiae* as a protein required for the "Start" event in the cell division cycle (11). In yeast, *cdc37* mutations cause synthetic lethality (as defined by Kaelin (56)) in combination with thermosensitive *cdc28* or *kin28* mutations (12, 13). The mammalian homolog of Cdc37, also referred to as $p50^{cdc37}$, was first identified within a complex consisting of the Rous sarcoma virus encoded oncogene $pp60^{v\text{-}scr}$ and Hsp90 (14, 15). Further studies confirmed that protein kinases are the most favored targets of Cdc37 (16). A plethora of other oncogenic kinases, including Raf isoforms, Akt, and ErbB2, have been identified as potential clients of Cdc37 (17–20). Expression of Cdc37 is up-regulated in cancer cells and tissues and can promote tumorigenesis when overexpressed (21, 22). In addition, Cdc37 can inhibit the ATPase activity of Hsp90 implied in promoting the assembly of kinase clients with both chaperone proteins (23, 24).

Structurally, the 44.5-kDa protein Cdc37 can be dissected into three domains (10, 25). Proteolytic fingerprinting studies indicate that it is comprised of an N-terminal domain (residues 1–127 ($Cdc37_N$), 15.5 kDa), a middle domain (residues 147–276 ($Cdc37_M$), 16 kDa), and a C-terminal domain (residues 283–378 ($Cdc37_C$), 10.5 kDa). The middle domain $Cdc37_M$ is

* This work was supported by the SPINE (Structural Proteomics in Europe) project of the European commission and by the Cluster of Excellence: Macromolecular Complexes (Deutsche Forschungsgemeinschaft). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 2W0G and 2K5B) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] Sanofi-Aventis Deutschland GmbH, Industriepark Höchst, Frankfurt am Main D-65926, Germany.

[2] To whom correspondence should be addressed: Tel.: 49-69-79829737; Fax: 49-69-79829515; E-mail: schwalbe@nmr.uni-frankfurt.de.

[3] The abbreviations used are: Hsp90, 90-kDa heat shock protein; Cdc37, cell division cycle protein 37; CSP, chemical shift perturbation; CST, cross-saturation transfer; HADDOCK, high ambiguity-driven biomolecular docking; HSQC, heteronuclear spin quantum correlation; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; RDC, residual dipolar couplings; r.m.s.d., root mean square deviation; R_g^{theo}, radius of gyration (theoretical); R_g^{exp}, radius of gyration (experimental); SAXS, small angle x-ray scattering; TROSY, transverse relaxation-optimized spectroscopy; DTT, dithiothreitol; PAA, polyacrylamide.

highly resistant to proteolytic digestion and was found to be the most stable domain of Cdc37 (26). Several studies (19, 27–30) have shown that $Cdc37_N$ binds to the client proteins (*e.g.* kinases). A highly conserved region within $Cdc37_N$ contains Ser-13 as a unique phosphorylation site for the protein kinase CKII. Phosphorylation seems to be an important mechanism for controlling the binding affinity to the client protein kinases and as a consequence thereof regulating the activity of multiple protein kinases (31). The Hsp90 binding site is located in the middle domain of Cdc37, which also comprises a putative dimerization region (32). An extended α-helix connects the Hsp90 binding site to the C-terminal domain, which might also be required for dimer formation (24, 32). The structure of the N-terminal domain is yet unknown but sequence analysis predicts a high content of α-helices (32).

The crystal structure of $Cdc37_{MC}$ (middle and C-terminal domain) bound to $Hsp90_N$ (from yeast) is a heterotetrameric complex, comprising Hsp90 bound to a Cdc37 dimer (32). The function of Cdc37 to inhibit the ATPase activity of Hsp90 (24) is due to the extended side chain of Arg-167 that binds to Glu-33 of Hsp90, thereby inhibiting the ATPase reaction. Additional studies indicate that $Cdc37_M$ exists as a monomer (26). Low resolution cryoelectron microscopy images of the Hsp90·Cdc37·Cdk4 complex show that Cdc37 is a monomer in this interaction (33). However, these studies utilized yeast Hsp90, and it was realized that it is necessary to investigate the interaction with human Hsp90 (26) to resolve the differences.

Here, we report the structure determination of the human $Hsp90_N$·$Cdc37_M$ complex based on the NMR backbone assignments of $Cdc37_M$ (34) and $Hsp90_N$ (35). The interface of the complex between human $Cdc37_M$ and $Hsp90_N$ could be mapped based on NMR chemical shift perturbation (CSP) studies, and cross-saturation transfer (CST) techniques. Subsequently, we solved the x-ray crystal structure of human $Cdc37_M$ at 1.88 Å in its monomeric form. The structure of the $Cdc37_M$·$Hsp90_N$ complex was generated by incorporating NMR data (CSPs and CSTs as ambiguous distance restraints and RDCs as orientational restraints) in the HADDOCK (36) docking protocol.

EXPERIMENTAL PROCEDURES

Cdc37 Cloning, Expression, and Purification—From the cDNA (Invitrogen) the base pairs corresponding to amino acids 1–126, 147–276, 1–283, and 1–378 of human Cdc37 were amplified by PCR using 5′-ACACAACCATGGTGGACTACAGCGTGTGGGACCAC-3′, 5′-ACACAAGGATCCTTCACTTGCTGAAGCCGTCTTTGCTGAGC-3′, 5′-AAACCATGGGGAAACACAAGACCTTCGTGGAAAAATACG-3′, 5′-AAGCAAGGATCCTTCACTTCATGGCCTTCTCGATGCGCAG-3′, 5′-ACACAAGGATCCTTCAGCGCTCCTCCTCCTCGTACTCCTTC-3′, 5′-TCTTTATTTTCAGGGCGCCATGGTGGACTACAGCGT-3′, and 5′-GCAGCCGGATCCTTCACACACTGACATCC-3′ as primers, respectively. The PCR product was double digested with NcoI and BamHI and cloned into the derivatized pKM263 (6xHis tag ProtGB1-Tev between Nde1 and Xho1) vector, which harbors ProtGB1 to enhance the solubility of the expressed fusion partner. The authenticity of the clone construct was confirmed by nucleotide sequencing. *Escherichia coli* strain BL21(DE3) (Novagen) containing the desired plasmid was grown in Luria-Bertani (LB) broth medium at 37 °C supplemented with 34 μg/ml chloramphenicol. Protein expression was induced with the addition of 1 mM isopropyl 1-thio-β-D-galactopyranoside when the A_{600} reached 0.6–0.7. The culture was incubated in 5-liter Erlenmeyer flask for 5–6 h with aeration (160 rpm) at 37 °C before the cells were harvested by centrifugation (5,000 × *g*, 15 min, 4 °C). The cell pellet was resuspended in lysis buffer (50 mM Tris-HCl (pH 8.0), 0.5 M NaCl, 5 mM of β-mercaptoethanol) with addition of 2 mg/ml lysozyme and 100 units/ml benzonase (Merck, Darmstadt, Germany) protein suspension. After sonication (10×, 60-s pulse on, and 180-s pulse off) in an ice bath, the cell debris was removed by centrifugation (18,000 × *g*, 45 min, 4 °C) to yield a supernatant containing the soluble protein. The supernatant was applied to a nickel-nitrilotriacetic acid FastFlow column (Qiagen) following the manufacturer's recommendations. After Tev-protease cleavage to remove the His tag ProtGB1, the protein was applied again on a nickel-nitrilotriacetic acid column. The pooled protein fractions were further purified by gel filtration on a Hiload 26/60 Superdex 75 preparative grade column (Amersham Pharmacia Biosciences) in Hepes 50 mM (pH 7.5), NaCl 100 mM, and 1 mM DTT. The protein-containing fractions were pooled, concentrated, and either immediately used for the experiments or stored at −20 °C. The ^{15}N-labeled protein was produced by using M9 minimal media with $^{15}NH_4Cl$ (1 g/liter) as the sole nitrogen source. Uniformly triple (^2H,^{13}C, and ^{15}N) or double (^2H and ^{15}N) labeled protein was produced using Silantes OD2 media. The degree of isotopic enrichment for the Silantes media was >98% for all isotopes (manufacturer's data). For selective labeling with ^{15}N-labeled amino acids, the following ^{15}N-labeled amino acids were used: alanine, arginine, lysine, methionine, isoleucine, leucine, valine, phenylalanine, tyrosine, and tryptophan (Spectra Stable Isotopes, Germany). The procedure applied for selective amino acid labeling is essentially the same as described in the literature (37).

$Hsp90_N$ Cloning, Expression, and Purification—From the cDNA (Stratagene) the base pairs corresponding to amino acids 18–223 (N-terminal domain) of human Hsp90 were amplified by PCR using 5′-AAACCATGGGGGAGACGTTCGCCTTTCAGG-3′ and 5′-GCAGCCGGATCCTCACTCCACAAAAAGAG-3′ as primers. The PCR product was double digested with NcoI and BamHI and was cloned into the derivatized pKM263, with a Protein G B1 domain (ProtGB1) fusion protein located C-terminally behind a stretch of six histidines and a tev cleavage site under the control of a T7-promoter (see above). The protein was expressed and purified in the same way as Cdc37, except for the post induction temperature, which was 25 °C. In addition, tetra-selective labeling with ^{15}N-glycine, ^{15}N-alanine, ^{15}N-phenylalanine, and ^{15}N-leucine was carried out as described above.

Due to the design of the expression vectors, Cdc37 and $Hsp90_N$ constructs carry two (GA) or four (GAMG) additional amino acids at the N terminus. High expression levels were achieved after induction with 1 mM isopropyl 1-thio-β-D-galactopyranoside at 37 °C for Cdc37 and 25 °C for $Hsp90_N$ and ~95% of the fused protein was found in the soluble fraction.

Typically, 30 mg/liter of pure (95%) proteins was obtained after cleavage of protein G. The final samples were dialyzed against the NMR buffer (Hepes 50 mM (pH 7.5), NaCl 100 mM, and 1 mM DTT).

Mutagenesis—Mutants of $Cdc37_M$ were created with QuikChange kit (Stratagene), and sequenced to confirm the incorporation of the correct mutation. The mutant proteins were purified following the same protocol as the wild-type $Cdc37_M$.

NMR Spectroscopy—NMR experiments were performed at 298 K, on Bruker 600-, 700-, 800-, 900-, and 950-MHz spectrometers equipped with cryogenic triple-resonance probes. For the $Cdc37_M$·$Hsp90_N$ complex a ratio of 1:1 was used for all measurements. The spectrometer was locked on D_2O. The spectra were processed using Topspin 2.0 (Bruker Biospin) and analyzed using either SPARKY 3.113[4] or CARA 1.8.3 (39). The resonances for the free form of $Cdc37_M$ at pH 5.5 and $Hsp90_N$ at pH 7.5 have been assigned before (34). Spectra of $Cdc37_M$ at pH 7.4 have been assigned by performing a pH series (5.0, 5.5, 6.0, 6.5, 7.0, 7.5, and 8.0) with 2,2-dimethyl-2-silapentane-5-sulfonic acid as an internal standard for calibration. Sequential backbone assignment of ^{15}N-labeled $Cdc37_M$ in complex with unlabeled $Hsp90_N$ and *vice versa* was achieved by overlaying the $^1H,^{15}N$-HSQC spectra of free $Cdc37_M$ (or $Hsp90_N$) and $Cdc37_M$ (or Hsp90) in complex and verified by the NOE patterns from a high resolution three-dimensional $^1H,^{15}N$-NOESY HSQC spectrum and selective ^{15}N-amino acid labeling (alanine, arginine, glycine, isoleucine, leucine, lysine, methionine, phenylalanine, tryptophan, tyrosine, and valine). Additionally, a three-dimensional $^1H,^{15}N$-NOESY HSQC spectrum for $Hsp90_N$ with a tetra-selectively labeled sample in which only glycine, alanine, phenylalanine, and leucine were ^{15}N-labeled was used to assist the assignment. $^1D(H,N)$ RDCs were measured for the $Cdc37_M$·$Hsp90_N$ complex (in which both proteins were ^{15}N-labeled) in pf1 phage (5 g/liter, Profos) and in 4% polyacrylamide (PAA) gel alignment media at 700 MHz. The $^1D(H,N)$ were extracted from IPAP-($^1H,^{15}N$)-HSQC spectra (40, 41). Signals that could be traced back reliably and determined unambiguously were analyzed using MODULE (42) and PALES (43). CST experiments were carried out at 25 °C according to the procedure of Takahashi *et al.* (57). The sample contained 1 mM of $^2H,^{15}N,^{13}C$-labeled $Cdc37_M$ protein with the unlabeled $Hsp90_N$ protein in $^1H_2O/^2H_2O$ (1:9, v/v) containing 50 mM Hepes (pH 7.5), 1 mM DTT, and 100 mM NaCl. The Gaussian pulse scheme was used for saturation of aliphatic protons with irradiation frequency set to 1.1 ppm. A saturation time of 2.0 s was employed.

X-ray Crystallography—Initial screening for crystallization conditions of $Cdc37_M$ was performed using screening kits purchased from JB (Jena Bioscience). These trials were successful, with five promising conditions producing crystals. The protein samples that did not give crystals were analyzed by mass spectrometry, which identified that they contained small fractions of proteins with an increased molecular mass of ~75, 150, and 228 Da. We attributed this observation to the formation of mixed disulfides of β-mercaptoethanol with the four free cysteines of $Cdc37_M$. Therefore, β-mercaptoethanol was omitted from the further protein preparations. The final diffracting quality crystals were grown from $Cdc37_M$ (Hepes, 50 mM, NaCl, 100 mM, pH 7.5, "protein buffer") at a final concentration of 1.1 mM, in solution containing Hepes, 100 mM, pH 7.5, sodium acetate, 100 mM, 22% polyethylene glycol 4000, ("reservoir buffer"). Crystal drops were setup using the sitting drop vapor diffusion method at 18 °C. Crystals were transferred in a stepwise manner into solutions of increasing glycerol concentrations of 5, 10, 15, 20, and 25%, respectively, based on one volume of protein buffer and two volumes of reservoir buffer, and flash-cooled in an X-Stream 2000 nitrogen stream at 140 K. Diffraction data were collected using a Rigaku R-Axis IV image plate detector at a wavelength of 1.5418 Å (CuKα radiation from a MicroMax-007 Microfocus generator). In addition to a high exposure data set consisting of 360 images of 1.0° rotations, a low exposure (one-third of the exposure time per image of that for high exposure) data set was collected consisting of 240 images of 1.5° rotations. The oscillation images were processed with the HKL/DENZO program package (44). The crystals belong to the hexagonal space group $P6_1$ with unit cell dimensions a = b = 48.67 Å and c = 104.26 Å. Scaling and merging of the oscillation images was performed with SCALEPACK (44). Data reduction and approximate absolute scaling of the data were performed with TRUNCATE from the CCP4 program suite (45). The phase problem was solved by molecular replacement using the program AMoRe (46) from the CCP4 suite, the diffraction data between 15 Å and 3 Å resolution, and a Cdc37 search model constructed from PDB entry 1US7 (32), extending from residue His-148 to residue Lys-276 and consisting of 1086 non-hydrogen atoms. The top solution with rotation angles $\alpha = 45.52°$, $\beta = 150.55°$, $\gamma = 311.52°$ and translations $T_x = -0.3117$ and $T_y = 0.0954$ (fractional coordinates) was associated with a correlation coefficient of 58.2% and an *R*-factor of 40.4%. After rigid body refinement with Refmac5 (47), the *R*-factor for the data between 42.14 Å and 1.88 Å was 43.8% ($R_{free}T = 43.9\%$). After several cycles of restrained refinement with Refmac5, the modeling of tightly bound water molecules, and manual rebuilding with COOT (48), the final structure was obtained.

Docking—Calculation of the $Cdc37_M$·$Hsp90_N$ complex was achieved using a high ambiguity driven docking (HADDOCK 2.0) approach with CNS 1.1 (49) essentially as described in the literature (36). The ambiguous interaction restraints have been defined for the residues that exhibited significant NMR amide chemical shift changes and/or strong saturation transfer upon interaction with its partner protein. For Hsp90, residues 117, 121, 123, 124, 125, 126, and 129 have been defined as active and residues 28, 32, 40, 46, 51, 54, 55, 114, 116, 119, 120, 122, 127, 130, 132, 133, and 134 as passive. The helix (112–124) and adjacent loop region from residues 106–127 were left fully flexible to offer freedom to the orientation of this helix. Additional hydrogen-bond restraints were added to keep the secondary structure of this helix. For $Cdc37_M$, residues 160, 161, 164, 165, 166, 167, 168, 193, 202, 204, 205, and 208 have been defined as active and residues 156, 157, 158, 169, 170, 174, 189, 192, 196, 197, 200, 201, 203, 207, 211, 242, and 249 as passive. The loop region 164–168 was left fully flexible during the docking. One

[4] T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco.

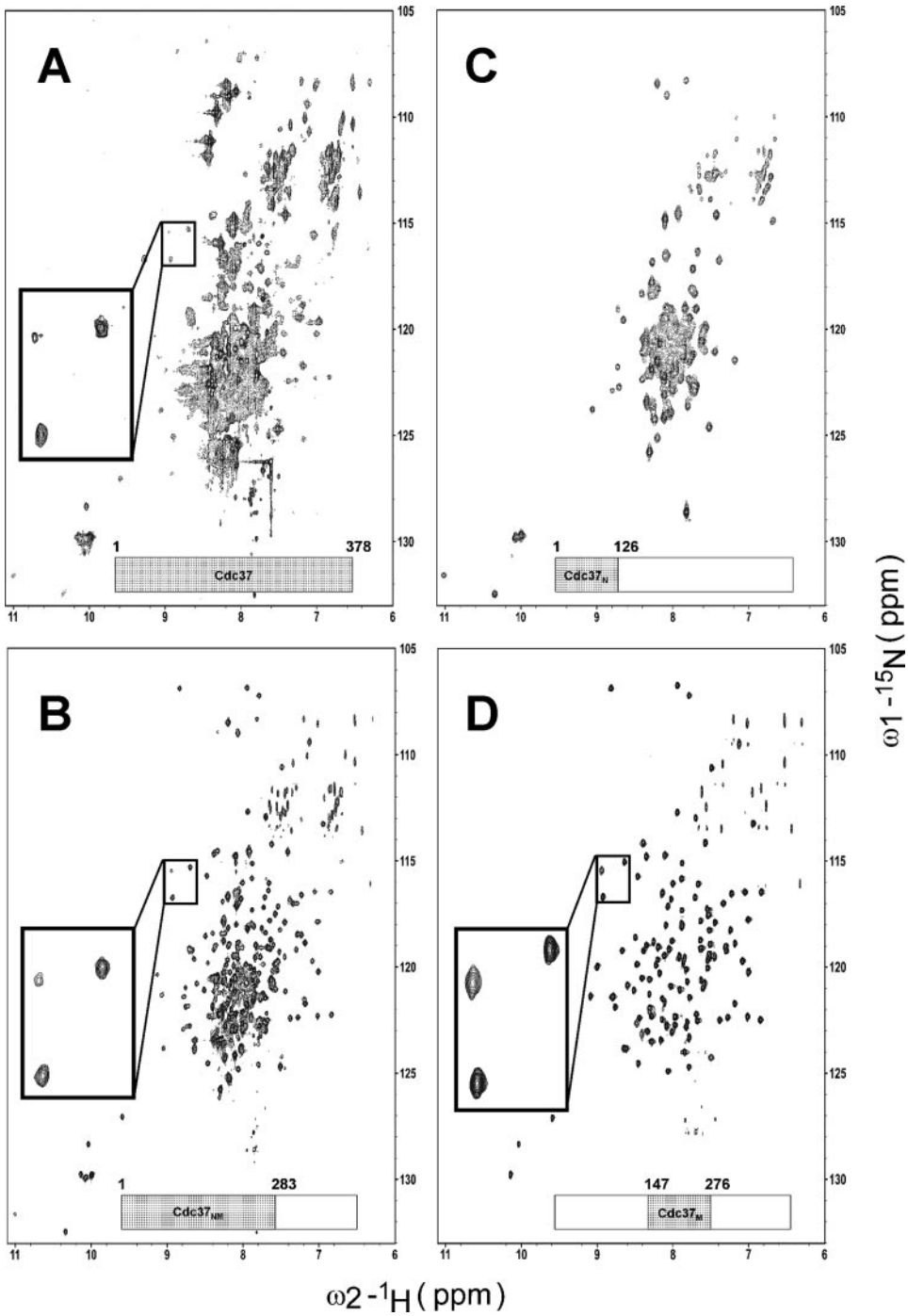


FIGURE 1. **^{1}H,^{15}N-TROSY spectra (950 MHz) of different Cdc37 constructs.** The boundaries of each construct (*e.g. A*, full-length; *B*, N-terminal and middle domain; *C*, N-terminal domain; and *D*, middle domain) are schematically indicated below each spectrum. The *insert* shows three peaks stemming from the middle domain ($Cdc37_M$) showing the same fold for this domain in all constructs. The spectra were acquired at 298 K in the same buffer solution (50 mM Hepes buffer (pH 7.5), 100 mM NaCl, 1 mM DTT, and 10% D_2O).

intermolecular NOE (Ala-117-HN to Ala-204-HB), as obtained from NOESY experiments, fitted very well with the initial docking results and was included in the final calculation. To enhance the convergence, ^{1}D(H,N) RDCs obtained both in 4% PAA gel and in 5 g/liter pf1 phage alignment media, were included in the calculation and the axial (D_a) and rhombic (D_r) components of the alignment tensor were estimated using PALES (43). For the initial rigid-body docking, only the RDCs in the stable secondary structure elements were used, whereas in the further refinement steps, all RDCs were used. The final values for D_a and D_r were −8.21 and 0.62 for the phage alignment and 9.39 and 0.45 for the alignment in the PAA gel. The dockings were performed using our $Cdc37_M$ crystal structure and the Hsp90 crystal structure (pdb: 1YES) (7). The protein allhdg 5.3 force field (50) was used for the calculations. The 200 final water-refined docking results are clustered at the interface within a threshold of 3.0 Å pairwise backbone r.m.s.d. The top-ranked ensemble, according to the average interaction energy and buried surface area, was accepted as the best representative of the complex.

RESULTS

Comparing Different Constructs of Cdc37 and Hsp90—The study of the rather large complex formed between the 45-kDa Cdc37 protein and the 23-kDa N-terminal domain of Hsp90 is a challenge for NMR spectroscopic studies. To study this complex, we reduced the boundaries of Cdc37 to obtain a construct that retains its structural and functional properties and is suitable for NMR studies. Beside the $Hsp90_N$ binding domain of Cdc37 (147–276, $Cdc37_M$), for which we recently published the NMR assignments (34), also the following Cdc37 constructs were investigated: full-length Cdc37 (1–378), the kinase binding domain (1–126, $Cdc37_N$), and both kinase and $Hsp90_N$ binding domains (1–283, $Cdc37_{NM}$). The suitability of the Cdc37 constructs for NMR was investigated by recording ^{1}H,^{15}N-TROSY spectra (Fig. 1). The spectrum of the full-length Cdc37 revealed considerable signal overlap and, in addition, signals that indicate unfolded regions (Fig. 1*A*). For $Cdc37_N$, the spectrum shows again signal overlap and evidence for a high content of partially unfolded and/or helical regions (Fig. 1*C*). $Cdc37_M$ shows a nicely resolved high resolution ^{1}H,^{15}N-TROSY spectrum (with 131 out of the 133 expected amide resonances), which indicates that this Hsp90-binding domain of Cdc37 is properly folded (34) (Fig. 1*D*). It was suspected that the $Cdc37_N$ might become more structured when it is together with the middle domain. Nevertheless, for

TABLE 1
Crystallographic data collection and refinement statistics

	All data	Highest resolution bin
Data		
Resolution limits (Å)	42.14-1.88	1.95-1.88
No. of measured reflections	368,290	41,732
No. of unique reflections	11,278	1,134
Redundancy	32.7	36.8
Completeness (%)	98.7	100
Completeness with I $>2\sigma$ (%)	94.1	86.1
I/σ(I)	49.8	18.7
R_{sym} (%)	5.8	30.6
B value from Wilson plot (Å^2)	35.5	
Mosaicity (°)	1.1	
Solvent content (%)	43	
Refinement		
Resolution limits (Å)	42.14-1.88	1.93-1.88
Test set selection	Random (5% = 609)	Random (5% = 46)
R_{free}T (%)	24.6	45.0
R_{work} (%)	21.2	29.1
$R_{work+test}$ (%)	21.3	
Number of non-hydrogen atoms	1130	
Mean B value (Å^2)	29.3	
r.m.s.d. values		
From ideal bond lengths (Å)	0.018	
From ideal bond angles (°)	1.544	
Ramachandran plot		
Favorable (%)	93.2	
Additional (%)	6.8	
Generous (number)	0	
Disallowed (number)	0	

the corresponding $Cdc37_{NM}$ construct the $^1H,^{15}N$-TROSY spectra indicated insufficient chemical shift dispersion (Fig. 1*B*). The amide signals originating from $Cdc37_M$ (shown specifically for three peaks as in the *insert* in Fig. 1*D*) have the same chemical shifts as in the full-length Cdc37 and $Cdc37_{NM}$ (Fig. 1, *A* and *B*), indicating that the fold of the $Cdc37_M$ is the same as in the full-length protein. Hence, the 16-kDa $Cdc37_M$ domain was used for all further NMR experiments. For Hsp90, the 23-kDa construct comprising the N-terminal ATP-binding domain (18–223) known to interact with Cdc37 was used throughout this study (35).

Structure of Human $Cdc37_M$—In the x-ray structure of the complex between yeast Hsp90 and human Cdc37, the middle domain of the Cdc37 protein forms a homodimeric interaction with a symmetry-related molecule (32). So far, the structure of Cdc37 is only available in the complex with Hsp90. However, by NMR, we do not see any evidence for dimer formation of $Cdc37_M$. To elucidate whether $Cdc37_M$ also forms dimers when crystallized without Hsp90, we solved the x-ray structure of free $Cdc37_M$. The obtained resolution after refinement was 1.88 Å with the characteristics summarized in Table 1. The structure of $Cdc37_M$ is entirely helical consisting of a 6-helix bundle (148–243) connected to a single long helix (245–276) as shown in Fig. 2. In the crystals, $Cdc37_M$ is a monomer, which is in excellent agreement with NMR diffusion-based measurements showing that $Cdc37_M$ exists as a monomer in solution with an $R_g^{exp} = 21.3$ Å and a theoretical (51) predicted $R_g^{theo} = 22.2$ Å. These findings agree very well with previous findings based on SAXS (26) and electron microscopy (33) data. The overall structure of this free-form monomeric $Cdc37_M$ is very similar to the x-ray structure of the dimeric $Cdc37_M$ when in complex with the yeast Hsp90 (32). The average r.m.s.d.

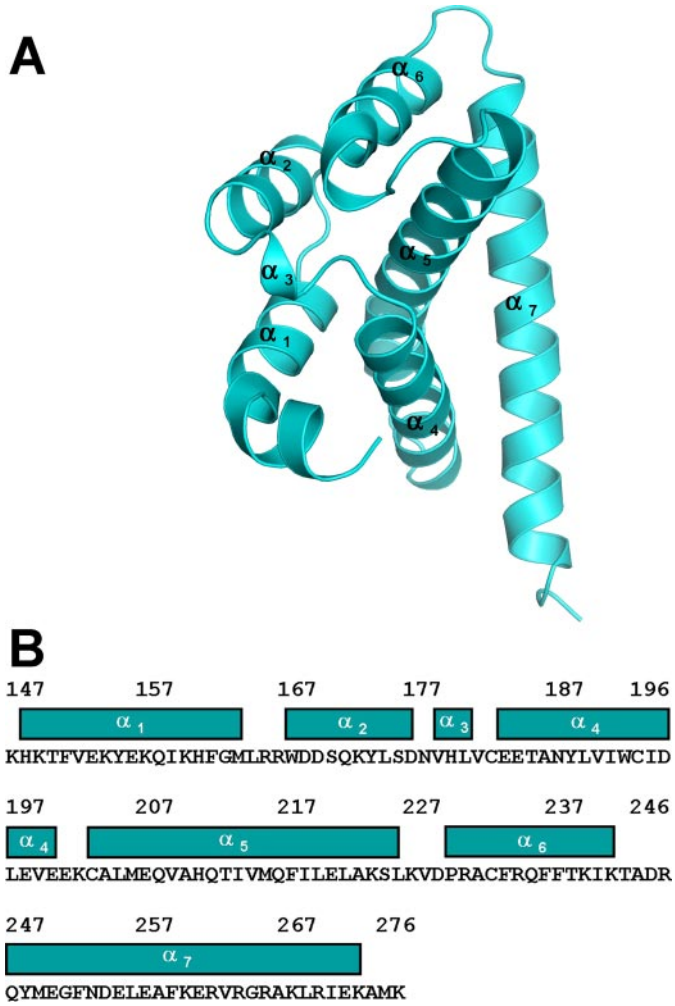


FIGURE 2. **X-ray structure of $Cdc37_M$.** *A*, x-ray structure of $Cdc37_M$ in *ribbon representation*. The helices are indicated. *B*, the amino acid sequence and the secondary structure boundaries of $Cdc37_M$ are represented schematically.

between the structures for all heavy atoms is 0.8 Å for the 6-helix bundle (148–243) and 1.3 Å for the full domain (148–276). The largest differences are observed in the loop connection (242–251) to the single long α_7 helix, which is exactly at the previously proposed dimerization interface (32). Interestingly, this region is also near to Hsp90 and may thus expose a potential additional interaction site. Furthermore, we found two possible conformations for the side chains of Arg-167 and Gln-208 in the $Cdc37_M$ crystals, which are also slightly different from the conformation found in the previously reported structure. Both residues reside in and near the interaction interface with Hsp90.

Assignment of Human $Cdc37_M$ and Human $Hsp90_N$—The $^1H,^{15}N$ and ^{13}C backbone and side-chain resonance assignment of the free-form $Cdc37_M$ has been reported previously at pH 5.5 (34). Because the interaction studies with $Hsp90_N$ have to be conducted at pH 7.5, we reassigned the amide resonances of $Cdc37_M$ at this different condition. Hence, we performed a series of NMR experiments ranging from pH 5.0 through pH 8.0 with a 0.5 pH unit increment, and used 2,2-dimethyl-2-silapentane-5-sulfonic acid as an internal standard for calibration. Using the $^1H,^{15}N$-HSQC spectra from the pH series we

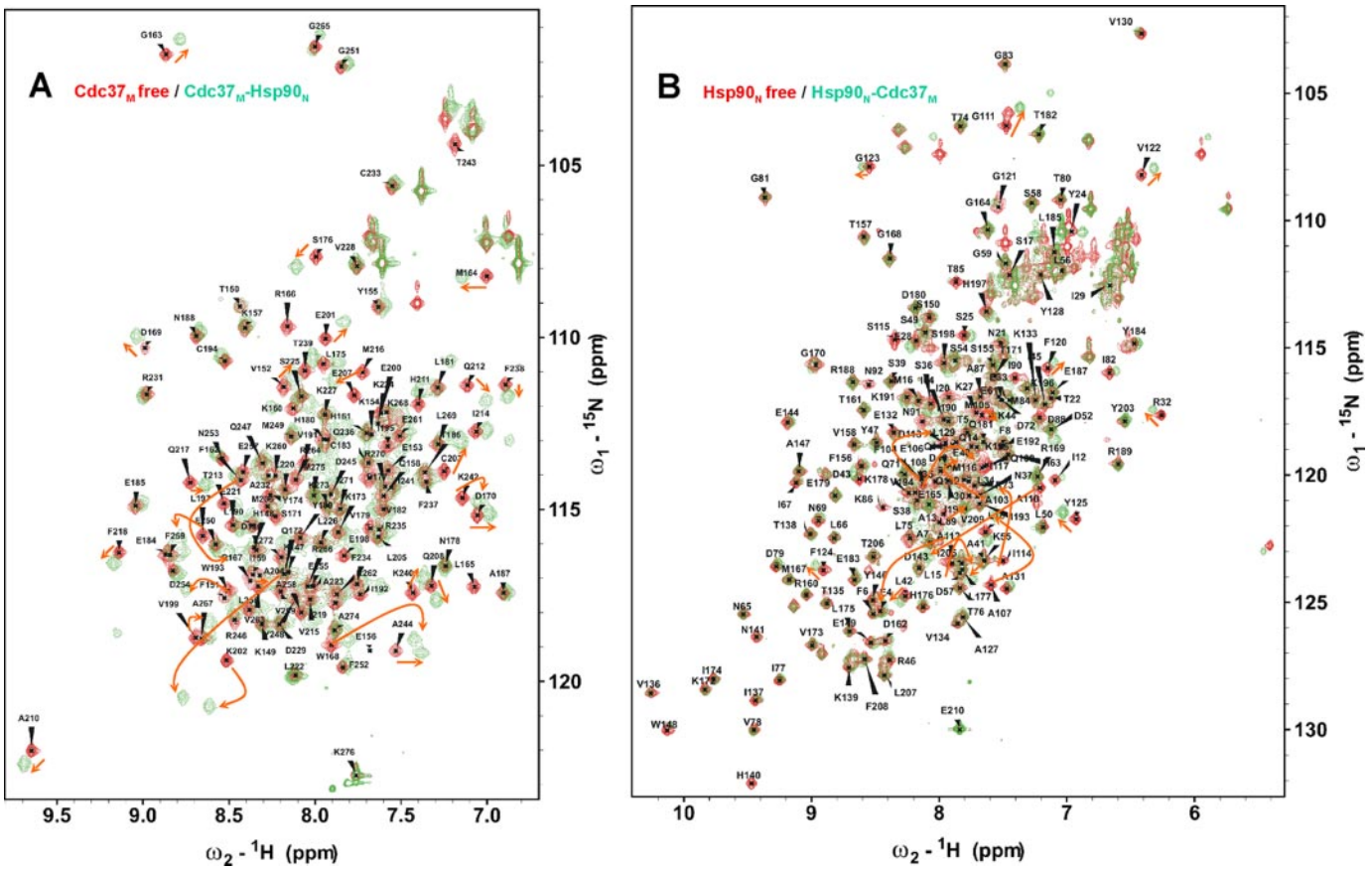


FIGURE 3. **$Cdc37_M$ binding to $Hsp90_N$.** Overlay of the ^1H,^{15}N-HSQC spectra (900 MHz)of $Cdc37_M$ in its free form(*red*) with $Hsp90_N$ bound form(*green*) (*A*) and of $Hsp90_N$ in its free (*red*) with $Cdc37_M$-bound form (*green*) (*B*). The backbone assignments are shown for the free form of $Cdc37_M$ (*A*) and $Hsp90_N$ (*B*), and some of the major peak shifts are indicated by *arrows*. The spectra were acquired at 298 K in the same buffer solution (50 mM Hepes buffer (pH 7.5), 100 mM NaCl, 1 mM DTT, and 10% D_2O).

were able to nearly completely assign $Cdc37_M$ (99% of the backbone amide resonances) at pH 7.5.

The ^1H,^{15}N and ^{13}C backbone and side-chain resonances of the human $Hsp90_N$ (18–223) at pH 7.5 in its free form have been assigned previously (35). The amide backbone assignment was complete with the exception of four stretches (Gln-23 to Glu-25, Leu-107 to Lys-116, Ser-165 to Gly-168, and Phe-213 to Tyr-216) and the amino acids Phe-37, Asn-40, Leu-45, Thr-65, Lys-74, Leu-76, Lys-84, Gly-132, Phe-138, Ser-140, Asp-156, Thr-176, Glu-199, and Gln-212.

Mapping of the Binding Interface of the $Cdc37_M$/$Hsp90_N$ Complex by NMR—To map the $Cdc37_M$ binding site for $Hsp90_N$, NMR titrations were performed with ^{15}N-labeled $Cdc37_M$ and unlabeled $Hsp90_N$ at 25 °C and pH 7.5. An overlay of the ^1H,^{15}N-HSQC spectra of free $Cdc37_M$ and the $Cdc37_M$·$Hsp90_N$ complex is shown in Fig. 3*A*. These spectra reveal significant changes in the combined amide chemical shifts (^1H,^{15}N) of the $Cdc37_M$ upon addition of $Hsp90_N$ (up to a molar ratio of 1:1). Many peaks in the two-dimensional ^1H,^{15}N-HSQC spectrum of $Cdc37_M$ are either significantly shifted or disappear (exchanged broadened) from the spectra indicating complex formation. The assignment of the shifted peaks was confirmed by using a combination of ^{15}N amino acid-selective labeling and three-dimensional ^1H,^{15}N-NOESY-HSQC spectra. To facilitate the assignment in the complex, $Cdc37_M$ was selectively labeled with ten different ^{15}N-labeled amino acids (Ala, Arg, Ile, Leu, Lys, Met, Phe, Trp, Tyr, and Val). Except for those signals that broadened beyond detection, all other peaks could be assigned unambiguously. The residues of $Cdc37_M$ revealing the largest combined (^1H,^{15}N) chemical shift perturbation (CSP $>$ 0.15 ppm) upon binding $Hsp90_N$ are Phe-162, Arg-167, Trp-168, Trp-193, Lys-202, Ala-204, Val-209, and Met-216 (Fig. 4*A*). Furthermore, amino acids Leu-165, Arg-166, Ile-195, Leu-205, and Glu-207 show a strong decrease in signal intensity. To further analyze the interaction site of $Cdc37_M$, we performed a CST experiment. This experiment was carried out using ^2H,^{15}N,^{13}C-labeled $Cdc37_M$ protein with the unlabeled $Hsp90_N$ protein in $^1H_2O/^2H_2O$ (1:9, v/v). The $Cdc37_M$ residues that appear to be in the interface ($<$0.78 intensity ratio) region include Lys-160, Met-164, Arg-166, Arg-167, Trp-168, Leu-205, Glu-207, Gln-208, and Ile-214 (Fig. 4*B*). When we combine CSPs, CSTs, and the decrease of signal intensity and map the binding site onto the $Cdc37_M$ crystal structure, two distinct regions appear to be involved in complex formation. The first region is composed of the loop (residues 164–167), and the second region is defined by the loop connecting α-helices 3 and 4 and the beginning region of the α-helix 4 (residues 203–208).

The reciprocal experiments using ^{15}N-labeled $Hsp90_N$ and unlabeled $Cdc37_M$ to identify the residues of human $Hsp90_N$ involved in the interaction with the $Cdc37_M$ were carried out as well. An overlay of the ^1H,^{15}N-HSQC spectra of free $Hsp90_N$

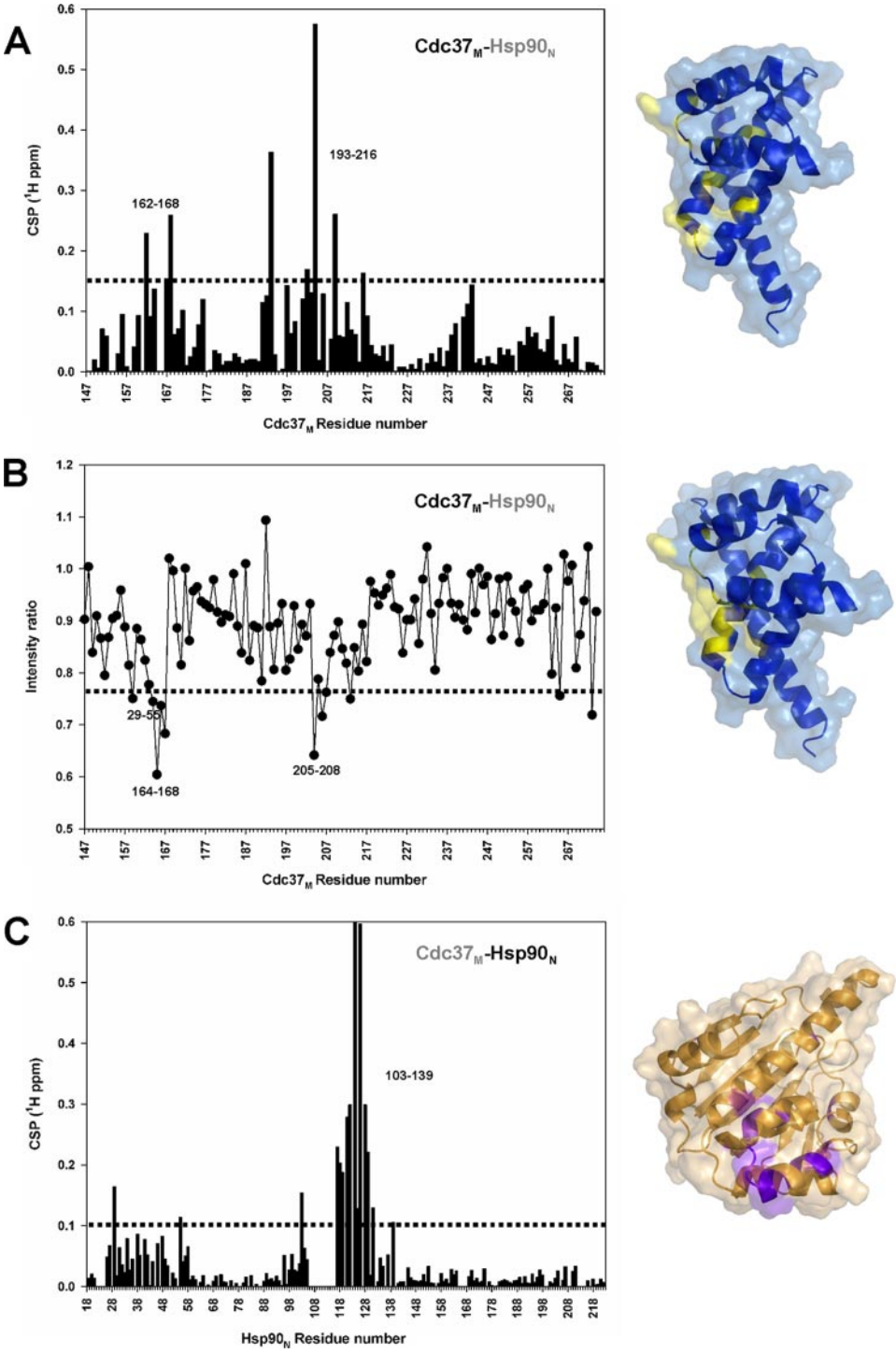


FIGURE 4. **Titration of $Cdc37_M$ with $Hsp90_N$ and *vice versa*.** *A*, correlation of CSPs *versus* residues in $Cdc37_M$ upon addition of $Hsp90_N$ (*left*) and surface representation of the interaction sites of $Cdc37_M$ (*yellow*) for $Hsp90_N$ (>0.15 ppm). *B*, plot of the intensity ratios of the cross-peaks in the cross-saturation transfer experiment (*left*). The intensity ratios of the backbone amide resonance of $Cdc37_M$ in complex with $Hsp90_N$ with irradiation to those of without irradiation measured in 10% H_2O/90% D_2O. Surface representation of $Cdc37_M$, where residues with intensity ratio of <0.78 are colored in *yellow* (*right*). *C*, correlation of CSPs *versus* residues in $Hsp90_N$ upon addition of $Cdc37_M$ (*left*) and surface representation of the interaction sites of $Hsp90_N$ (*purple*) for $Cdc37_M$ (>0.1 ppm). For the CSPs, amide ^1H and ^{15}N resonance shifts have been mapped and combined as Euclidian distances between peak maxima taking into account the gyromagnetic ratio of proton and nitrogen (in ^1H ppm). *Missing bars* indicate prolines or amide resonances that have not been assigned.

and the $Hsp90_N$·$Cdc37_M$ (molar ratio 1:1) complex is shown in Fig. 3*B*. The shifted amide resonances of $Hsp90_N$ were assigned using a combination of selectively labeled samples and three-dimensional ^1H,^{15}N-NOESY-HSQC spectrum. $Hsp90_N$ was selectively labeled with six different ^{15}N-labeled amino acids (Ala, Gly, Ile, Leu, Met, and Phe). An additional three-dimensional ^1H,^{15}N-NOESY-HSQC spectrum was recorded for a sample in which four different amino acids (Ala, Gly, Phe, and Leu) were selectively ^{15}N-labeled. The $Hsp90_N$ amide CSPs upon binding to $Cdc37_M$ were analyzed to characterize the amino acids potentially involved in the interaction. The largest combined (^1H,^{15}N) chemical shift changes (>0.10 ppm) are observed for a few residues in the first, second, and fifth α-helix (residues 29, 55, and 103, respectively), and many residues in the sixth α-helix (residues 117–119 and 121–126) as well as in the seventh α-helix (residues 128, 129, and 131). These residues define two distinct $Hsp90_N$ interacting sites for $Cdc37_M$, which are located on one side of the molecule and are localized around a hydrophobic patch (Fig. 4*C*).

Determination of the Human $Cdc37_M$·$Hsp90_N$ Complex Structure—Based on the interaction interface as determined from CSPs and CSTs (Fig. 5, *A* and *B*) we calculated the structure of the binary $Cdc37_M$·$Hsp90_N$ complex using HADDOCK (36). The calculations were performed using the crystal structures of human $Cdc37_M$ determined here and $Hsp90_N$ in the open conformation (pdb: 1YES (7)). The initial docking model proposed a number of possible solutions that differed in the relative orientation of the proteins (Fig. 5*C*). To resolve this ambiguity, we measured two sets of RDCs of the complex. RDCs could be obtained using either 5 g/liter phage pf1 phage alignment medium (Profos) or alignment in 4% polyacrylamide gels. Both sets of ^1D(N,H) RDC values were extracted from IPAP-(^1H,^{15}N)-HSQC spectra and yielded a linearly independent overall alignment tensor (Fig. 5*C*) and thus provided additional information on the relative orientation of the two proteins. The RDCs have been measured on a sample containing both ^{15}N-labeled $Cdc37_M$ and ^{15}N-labeled $Hsp90_N$ in a 1:1 ratio.

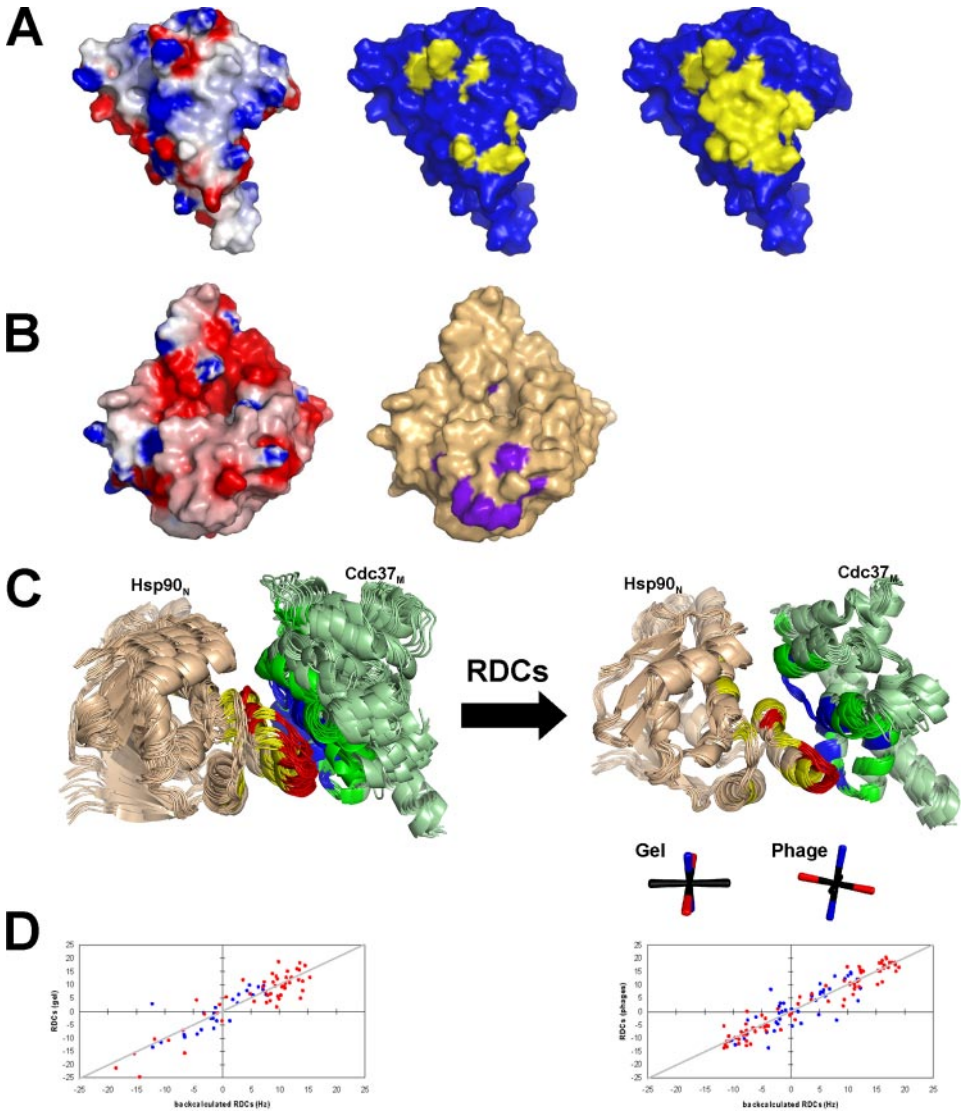


FIGURE 5. **NMR data driven docking model of human $Cdc37_M$·$Hsp90_N$ complex.** *A*, surface representation of $Cdc37_M$. *Left*: The molecular surface is *color-coded* by electrostatic potential, as calculated with DELPHI (38). Potentials less than −10 kT are *red*, those greater than +10 kT are *blue*, and neutral potentials (0 kT) are *white*. The *color* in between is produced by linear extrapolation. *Middle*, CSPs from Fig. 4*A* (>0.15 ppm) in *yellow* mapped onto the structure. *Right*, cross-saturation transfer experiment (Fig. 4*B*) in which residues with intensity ratios <0.78 are mapped onto the surface (*yellow*). *B*, surface representation of human $Hsp90_N$ (1YES.pdb). *Left*, the molecular surface is *color-coded* by electrostatic potential, as calculated with DELPHI (38). *Right*, CSPs from Fig. 4*C* (>0.1 ppm) in *purple* mapped onto the structure. *C*, bundle of the $Cdc37_M$·$Hsp90_N$ complex as obtained using HADDOCK, with and without using RDCs. For the calculation without RDCs (*left*), the ten best docking results for each of the two major HADDOCK clusters are shown, and for the calculation with RDCs (*right*) the ten best docking results are shown. The active/passive restraints were used as defined under "Experimental Procedures" and are *colored*, respectively, in *red* and *yellow* for $Hsp90_N$ or *blue* and *dark-green* for $Cdc37_M$. The RDC tensors for alignment in either polyacrylamide (PAA) gel or pf1 Phages are indicated below. *D*, correlation of the experimental *versus* back calculated RDCs for the free-form x-ray structures of $Cdc37_M$ (*blue*) and $Hsp90_N$ (*red*). Protons were added to the x-ray structures using CNS 1.1 (49) without using the RDCs. RDCs were back-calculated using the program PALES (43) with the bestFit option. *Left*: correlation plot in PAA gel alignment medium with Cdc37 in *blue* (25 RDCs, correlation coefficient of 0.822) and $Hsp90_N$ in *red* (52 RDCs, correlation coefficient of 0.903). *Right*: correlation plot in pf1 phages alignment medium with Cdc37 in *blue* (42 RDCs, correlation coefficient of 0.816) and $Hsp90_N$ in *red* (79 RDCs, correlation coefficient of 0.964).

Although some amide signals of both proteins overlap, there is sufficient dispersion to identify many residues from $Cdc37_M$ or $Hsp90_N$. In the phage alignment media, we could reliably identify 42 ^{1}D(N,H) RDCs for $Cdc37_M$ and 79 ^{1}D(N,H) RDCs for $Hsp90_N$. Despite the line broadening in the polyacrylamide gel alignment, we could reliably identify 25 ^{1}D(N,H) RDCs for $Cdc37_M$ and 52 ^{1}D(N,H) RDCs for $Hsp90_N$. The RDCs fit already very well with the free-form protein structures of $Cdc37_M$ and $Hsp90_N$ (Fig. 5*D*) indicating that their overall structure will not change much when in complex. In addition to the RDCs, we measured intermolecular NOEs. Two-dimensional ^{15}N-edited NOESY spectra were recorded for the free ^{15}N,^{2}H $Cdc37_M$ and $Hsp90_N$ proteins as well as two-dimensional and three-dimensional ^{15}N-filtered NOESY spectra for the ^{15}N,^{2}H-labeled protein when in complex with the ^{14}N,^{1}H (unlabeled) partner. In these experiments, we could identify several intermolecular cross-peaks. However, because no full side-chain assignment is available, in most cases we could not unambiguously assign them. Based on dockings using the interaction interface either with or without the RDCs, we included one strong intermolecular NOE as observed for $Hsp90_N$ Ala-117-HN to $Cdc37_M$ Ala-204-HB in the final docking calculation.

In the final docking calculations, the active and passive residues in the interaction site and the flexible interfaces have been included as stated in under "Experimental Procedures" as well as both sets of RDCs and the strong intermolecular NOE. The top-ranked docking model for the $Cdc37_M$·$Hsp90_N$ complex (Fig. 5*C*) shows the preferential binding mode and includes 185 out of 200 final structures when clustered within the 3.0 Å pairwise backbone r.m.s.d. These final structures have been analyzed whether they satisfy the biochemical and structural restraints. The active residue 168 of $Cdc37_M$ causes violations at a 2 Å ambiguous interaction restraint cut-off, but is still acceptable at a distance of 3 Å from $Hsp90_N$.

$Cdc37_M$-$Hsp90_N$ Interface and Structural Reorganization upon Complex Formation—The interface between $Cdc37_M$ and $Hsp90_N$ buries a total surface area of 1579 ± 92 Å^2, which is highly comparable to the interface for the average protein·protein complex (1600 ± 400 Å^2) (52). There are two main $Cdc37_M$·$Hsp90_N$ interaction regions that together form the interface. The hydrophobic core is mainly formed by the interaction of $Cdc37_M$ residues in the α_4 and α_5 helix with the residues in the $Hsp90_N$ α_6 helix and adjacent

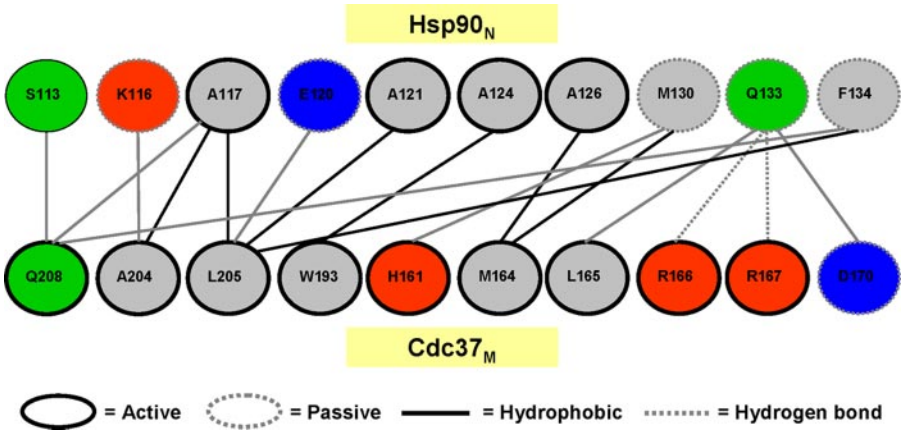


FIGURE 6. **Schematic representation of the interaction interface of NMR calculated $Cdc37_M$·$Hsp90_N$ complex structure.** The positively charged, negatively charged, neutral, and hydrophobic amino acids are colored in *red*, *blue*, *green*, and *gray*, respectively.

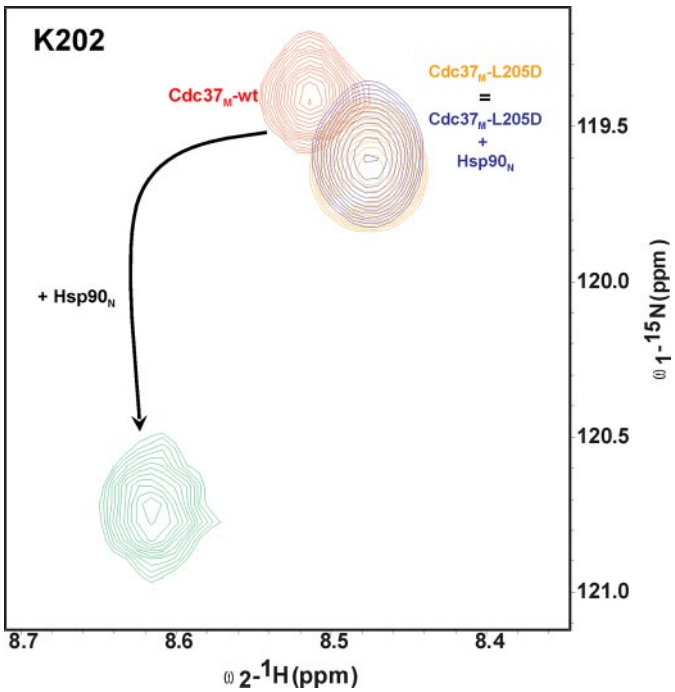


FIGURE 7. **$Cdc37_M$-L205D mutant does not bind to $Hsp90_N$.** ^1H,^{15}N-HSQC spectral region showing the amide signal for Lys-202, for wild type-free (*red*), mutant-free (*orange*), wild type in complex with $Hsp90_N$ (1:1) (*green*), and mutant in the presence of $Hsp90_N$ (1:4) (*blue*).

residues. Additionally, the $Cdc37_M$ α_1, α_2 helices and connecting loop interact with the $Hsp90_N$ α_7 helix and adjacent loop regions, involving many charged residues. Our docking results show an ensemble with many electrostatic and hydrophobic contacts as shown in Fig. 6. Electrostatic interactions are generally observed between the residues Arg-166, Arg-167, and Asp-170 of $Cdc37_M$ and Gln-133 of $Hsp90_N$ (hydrogen bonds are often observed for $Cdc37_M$ residues Arg-166 and Arg-167 to residue Gln-133 of $Hsp90_N$). The hydrophobic core involves residues Met-164, Trp-193, Ala-204, and Leu-205 of $Cdc37_M$ and Ala-117, Ala-121, Ala-124, Ala-126, Met-130, and Phe-134 of $Hsp90_N$. Upon $Hsp90_N$ binding, there are no large conformational changes in $Cdc37_M$, except in the flexible loop regions of the molecule. When superimposing the $Cdc37_M$·$Hsp90_N$ complex onto the uncomplexed molecules (coordinates from our $Cdc37_M$ crystal structure and 1YES for $Hsp90_N$), the r.m.s.d. on the C_α atoms is 0.7 Å for $Cdc37_M$ and 0.8 Å for $Hsp90_N$ (0.5 Å for $Hsp90_N$ when excluding residues 106–127). The difference in the helical region (residues 106–127) of $Hsp90_N$ is due to the involvement of this region as a hydrophobic core in the complex structure.

Residue Leu-205 of Cdc37 Is Critical for the Complex Formation—Our structure of the $Hsp90_N$·$Cdc37_M$ complex indicates the important residues involved at the protein-protein interface. These residues can be divided into two categories: one forms typical salt bridge interactions, and the other forms the hydrophobic core of the complex as described previously. NMR cross-saturation transfer data show that Leu-205 in $Cdc37_M$ experiences the largest decrease in signal intensity. To assess the role of Leu-205, we mutated this residue to phenylalanine, isoleucine, alanine, glycine, aspartic acid, and glutamic acid. The two-dimensional ^1H,^{15}N-HSQC spectra of all mutants typically retain the spectra of the wild-type, except for the region around the mutation, indicating that there are no large conformational changes brought about by the mutation. Typically, NMR can identify protein-protein interactions even with as low affinity as ~5 mM (K_D). The wild-type $Cdc37_M$ binds to $Hsp90_N$ with a K_D of 4.0 μM (26, 32) and shows strong changes in the ^1H,^{15}N-HSQC spectrum upon complex formation. For mutants L205D and L205Q, changes in the ^1H,^{15}N-HSQC cannot be observed ([$Cdc37_M$]:[$Hsp90_N$] = 1:4), indicating that those $Cdc37_M$ mutants do not bind anymore to $Hsp90_N$ (Fig. 7). Control experiments showed that the L205F and L205I mutants retained the interaction, whereas there was a significant decrease in the affinity for mutants L205A and L205G. The observation supports that residue Leu-205 of $Cdc37_M$ drives the complex formation by its hydrophobic nature.

Comparison with Previously Published Yeast $Hsp90_N$-Human Cdc37 Crystal Structure—In our study we used human $Hsp90_N$, which has 69% sequence identity (Fig. 8) with that of the yeast homologue (32). However, the residues of Hsp90 in contact with Cdc37 are well conserved to a large extent, with the exception of Gln-123, which is serine in the case of yeast homologue. A superposition of our structure of the human $Cdc37_M$·$Hsp90_N$ complex with the Cdc37 (human)-Hsp90 (yeast) crystal structure (pdb: 1US7) (32) reveals a similar overall conformation, with an r.m.s.d. of 1.0 Å, when fitted onto the C_α atoms of the Cdc37. The interface of our NMR-based solution structure has a buried surface area of ~1600 Å^2, which is distinctly larger than the ~1100 Å^2 of the crystal structure. Our human $Cdc37_M$·$Hsp90_N$ structure is thus more compact than the previously reported heterogeneous structure. The most

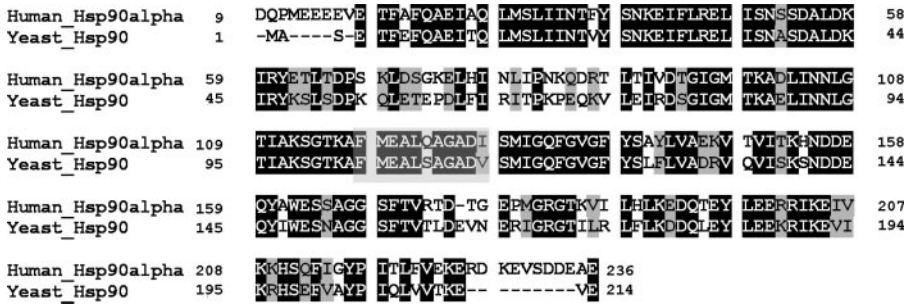


FIGURE 8. **Sequence alignment of human Hsp90$_N$ and yeast Hsp90$_N$.** The *shaded region* is the region that interacts with Cdc37$_M$ and forms the hydrophobic core of the interface.

eminent differences in Cdc37$_M$ are observed in the α_6-α_7-loop region. This difference already observed in the free-form crystal structure and very likely to arise from the fact that Cdc37 exists as a dimer in the crystal structure. Furthermore, changes are observed for the orientation of some side chains (Arg-167 and Gln-208) in the α_1-α_2 region of Cdc37$_M$.

DISCUSSION

Studying the protein-protein interaction of the full-length human Cdc37 (~45 kDa) with human Hsp90$_N$ (~23 kDa) at molecular detail by NMR is most challenging due to the relative large size of this complex. Because the best quality NMR spectra was obtained with the Cdc37$_M$ domain, which is known to interact with Hsp90, this study was focused on the examination of the Cdc37$_M$ structure and its interaction with Hsp90$_N$. An x-ray structure of human Cdc37$_M$ in complex with yeast Hsp90$_N$ was already available (32). In this study, an albeit longer Cdc37 construct (amino acids 148–347) was used. Interestingly, this construct was found to form a dimer. The core of the dimer interface has been reported to be around Gln-247 and Tyr-248 and the prolonged C terminus, which forms a small three-helix bundle (292–347), is not at all involved in dimer formation (32). However, we did not observe any evidence for dimer formation with the constructs used in our study. Also our experiments on the Cdc37$_M$·Hsp90$_N$ complex did not show signals that indicate formation of larger complexes. The existence of Cdc37$_M$ as a monomer was strongly supported by the high resolution x-ray structure of uncomplexed Cdc37$_M$ (Fig. 2). It should be noted, however, that the overall structure of monomeric and dimeric Cdc37$_M$ remains the same. The largest differences between the structures are found in the previously reported dimer interface, which is most probably caused by crystal packing and which is interestingly near to Hsp90 in the complex. Our findings for monomeric Cdc37 are also consistent with recent observations from Vaughan *et al.* (33) who analyzed a complex of Hsp90·Cdc37·Cdk4 by cryoelectron microscopy.

The interaction interface between human Cdc37$_M$ and Hsp90$_N$ was mapped based on NMR studies. Mapping of the CSPs and CSTs onto the structures reveals two eminent surface areas corresponding to a small area with charged residues and a large hydrophobic area. Subsequently, the docking program HADDOCK was used to calculate the NMR-based structure of the complex. When only using the ambiguous distance restraints that are based on all the existing data (CSPs, CSTs, and NOE) for the docking calculation, the final solutions did not converge and yielded two possible clusters with a different orientation. A single solution was obtained when we incorporated RDCs in our calculations. The overall conformation of the final NMR-based structure of the human complex is very similar to the known crystal structure of the heterogeneous complex. However, there are some clear differences observed for specific intermolecular contacts in the distinctly larger interaction interface, most probably causing this structure to be slightly more compact.

The Hsp90 residues in the interface of the complex are very well conserved, except for Gln-123, which is serine in case of the yeast homologue. In the crystal structure with the yeast homologue of Hsp90, this serine residue forms an intermolecular hydrogen bond with the Lys-202 side chain of Cdc37. In our model of the human complex, however, we find that the Gln-123 side chain is just outside the binding interface, and makes no significant contribution to the interaction. Instead, in this case, the nearby side chain of the Glu-120 residue of the human Hsp90 generally forms a hydrogen bond with the Lys-202 side chain of Cdc37. An important hydrogen bond is formed between Glu-47 (human sequence numbering) of Hsp90 and Arg-167 of Cdc37, which is considered to be the key interaction that arrests the ATPase action of Hsp90 (32). In the best structure from our docking bundle, we indeed also observe this hydrogen bond. However, we do not observe this hydrogen bond in all the docking structures, which is most probably a result of the two possible conformations for the side chain of Arg-167 that we observed in our free-form Cdc37$_M$ x-ray structure, that we use for our docking. We suppose that this side chain is flexible in the free-form Cdc37 and stabilizes the complex by forming an intermolecular hydrogen bond with Hsp90. Furthermore, in the heterogeneous crystal structure of the complex, the backbone oxygen of Hsp90 residue Gln-133 (human sequence numbering) accepts an intermolecular hydrogen bond from the side-chain Gln-208 of Cdc37, and its side-chain oxygen forms intermolecular hydrogen bonds with the backbone Arg-166 and Arg-167 amides of Cdc37. In our human complex we did not observe the interaction between the Hsp90 residue Gln-133 and the Cdc37 residue Gln-208. This could again be a result of the two possible conformations observed for this Gln-208 side chain we observed in the free-form x-ray structure, but on the other hand for the different orientation of this side chain in our human model, we observed that the amino moiety is rather close to the side-chain hydroxy of the Ser-113 residue of Hsp90, with which it may possibly form an intermolecular hydrogen bond. The hydrogen bonds between the side chain of Hsp90 residue Gln-133 and the backbone Arg-166 and Arg-167 amides of Cdc37 is very well maintained in our human structure of the complex. Furthermore, the hydrophobic core in our NMR-based structure of the com-

plex includes the proximity of Ala-124 of $Hsp90_N$ and Trp-193 of $Cdc37_M$, which are slightly further apart from each other in the heterogeneous crystal structure.

In recent years, the Hsp90 protein became attractive as a pharmaceutical anti-cancer target. Many protein kinases, of which some might be causative for cancer when deregulated or overexpressed, require Hsp90 to be maintained in a functional state. Targeting of Hsp90 thus seems to be attractive, because several protein kinases can be targeted at the same time. The Hsp90-specific inhibitor, 17-allylamino-17-demethoxy geldanamycin is already undergoing clinical trials (53). Hsp90-specific inhibitors known so far are directed against the ATP-binding pocket of Hsp90. Because Hsp90 needs co-chaperones to fulfill its function and one of these, Cdc37, is mainly found in combination with regulation of protein kinases, the whole Hsp90·Cdc37 complex could be considered as an interesting drug target. In addition, increased levels of Cdc37 have shown to enhance cancer cell growth (54). A drug that inhibits the formation of this complex might be a very valuable compound. Knowing the interaction interface of the Cdc37·Hsp90 is a prerequisite for designing substances inhibiting formation of the Hsp90·Cdc37. In our study we analyzed the human $Cdc37_M$·$Hsp90_N$ complex in great detail, and identified Leu-205 of Cdc37 as a key residue. This residue resides in the interface and makes many contacts with different residues of Hsp90 to stabilize the complex. Mutational analysis of Leu-205 indicates that the hydrophobic nature of this critical residue promotes the complex formation.

Very recently, a compound was identified that disrupts the Hsp90·Cdc37 complex in cancer cells (55) supporting the idea to target the Hsp90-Cdc37 interaction. We expect that the structure of the human $Cdc37_M$·$Hsp90_N$ complex solved in this study can be used to assist further studies in the development of small molecule inhibitors.

REFERENCES

1. Ferrarini, M., Heltai, S., Zocchi, M. R., and Rugarli, C. (1992) *Int. J. Cancer* **51,** 613–619
2. Yufu, Y., Nishimura, J., and Nawata, H. (1992) *Leuk. Res.* **16,** 597–605
3. Welch, W. J., and Feramisco, J. R. (1982) *J. Biol. Chem.* **257,** 14949–14959
4. Picard, D. (2002) *Cell. Mol. Life Sci.* **59,** 1640–1648
5. Terasawa, K., Minami, M., and Minami, Y. (2005) *J. Biochem.* (*Tokyo*) **137,** 443–447
6. Wegele, H., Mueller, L., and Buchner, J. (2004) *Rev. Physiol. Biochem. Pharmacol.* **151,** 1–44
7. Stebbins, C. E., Russo, A. A., Schneider, C., Rosen, N., Hartl, F. U., and Pavletich, N. P. (1997) *Cell* **89,** 239–250
8. Whitesell, L., and Lindquist, S. L. (2005) *Nat. Rev. Cancer* **5,** 761–772
9. Hunter, T., and Poon, R. Y. C. (1997) *Trends Cell Biol.* **7,** 157–161
10. MacLean, M., and Picard, D. (2003) *Cell Stress Chaperones* **8,** 114–119
11. Reed, S. I. (1980) *Genetics* **95,** 561–577
12. Reed, S. I., de Barros Lopes, M. A., Ferguson, J., Hadwiger, J. A., Ho, J. Y., Horwitz, R., Jones, C. A., Lorincz, A. T., Mendenhall, M. D., Peterson, T. A., Richardson, S. L., and Wittenberg, C. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50,** 627–634
13. Valay, J. G., Simon, M., Dubois, M. F., Bensaude, O., Facca, C., and Faye, G. (1995) *J. Mol. Biol.* **249,** 535–544
14. Whitelaw, M. L., Hutchison, K., and Perdew, G. H. (1991) *J. Biol. Chem.* **266,** 16436–16440
15. Brugge, J. S. (1986) *Curr. Top. Microbiol. Immunol.* **123,** 1–22
16. Mort-Bontemps-Soret, M., Facca, C., and Faye, G. (2002) *Mol. Genet. Genomics* **267,** 447–458
17. Basso, A. D., Solit, D. B., Chiosis, G., Giri, B., Tsichlis, P., and Rosen, N. (2002) *J. Biol. Chem.* **277,** 39858–39866
18. Lamphere, L., Fiore, F., Xu, X., Brizuela, L., Keezer, S., Sardet, C., Draetta, G. F., and Gyuris, J. (1997) *Oncogene* **14,** 1999–2004
19. Grammatikakis, N., Lin, J. H., Grammatikakis, A., Tsichlis, P. N., and Cochran, B. H. (1999) *Mol. Cell. Biol.* **19,** 1661–1672
20. Pearl, L. H. (2005) *Curr. Opin. Genet. Dev.* **15,** 55–61
21. Schwarze, S. R., Fu, V. X., and Jarrard, D. F. (2003) *Cancer Res.* **63,** 4614–4619
22. Stepanova, L., Yang, G., DeMayo, F., Wheeler, T. M., Finegold, M., Thompson, T. C., and Harper, J. W. (2000) *Oncogene* **19,** 2186–2193
23. Matts, R. L., and Caplan, A. J. (2007) in *Heat Shock Proteins in Cancer* (Calderwood, S. K., Sherman, M. Y., and Ciocca, D. R. eds.) 1st Ed., pp. 331–350, Springer Publishers, Dordrecht, The Netherlands
24. Siligardi, G., Panaretou, B., Meyerc, P., Singh, S., Woolfson, D. N., Piper, P. W., Pearl, L. H., and Prodromou, C. (2002) *J. Biol. Chem.* **277,** 20151–20159
25. Shao, J., Irwin, A., Hartson, S. D., and Matts, R. L. (2003) *Biochemistry* **42,** 12577–12588
26. Zhang, W., Hirshberg, M., McLaughlin, S. H., Lazar, G. A., Grossmann, J. G., Nielsen, P. R., Sobott, F., Robinson, C. V., Jackson, S. E., and Laue, E. D. (2004) *J. Mol. Biol.* **340,** 891–907
27. Lee, P., Rao, J., Fliss, A., Yang, E., Garrett, S., and Caplan, A. J. (2002) *J. Cell Biol.* **159,** 1051–1059
28. Scholz, G., Hartson, S. D., Cartledge, K., Hall, N., Shao, J., Dunn, A. R., and Matts, R. L. (2000) *Mol. Cell. Biol.* **20,** 6984–6995
29. Shao, J., Grammatikakis, N., Scroggins, B. T., Uma, S., Huang, W., Chen, J. J., Hartson, S. D., and Matts, R. L. (2001) *J. Biol. Chem.* **276,** 206–214
30. Shao, J., Prince, T., Hartson, S. D., and Matts, R. L. (2003) *J. Biol. Chem.* **278,** 38117–38120
31. Miyata, Y., and Nishida, E. (2004) *Mol. Cell. Biol.* **24,** 4065–4074
32. Roe, S. M., Ali, M. M. U., Meyer, P., Vaughan, C. K., Panaretou, B., Piper, P. W., Prodromou, C., and Pearl, L. H. (2004) *Cell* **116,** 87–98
33. Vaughan, C. K., Gohlke, U., Sobott, F., Good, V. M., Ali, M. M. U., Prodromou, C., Robinson, C. V., Saibil, H. R., and Pearl, L. H. (2006) *Mol. Cell* **23,** 697–707
34. Sreeramulu, S., Kumar, J., Richter, C., Vogtherr, M., Saxena, K., Langer, T., and Schwalbe, H. (2005) *J. Biomol. NMR* **32,** 262
35. Jacobs, D. M., Langer, T., Elshorst, B., Saxena, K., Fiebig, K. M., Vogtherr, M., and Schwalbe, H. (2006) *J. Biomol. NMR* **36,** 52
36. Dominguez, C., Boelens, R., and Bonvin, A. M. J. J. (2003) *J. Am. Chem. Soc.* **125,** 1731–1737
37. Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., and Dahlquist, F. W. (1989) *Methods Enzymol.* **177,** 44–73
38. Honig, B., Sharp, K., and Yang, A. S. (1993) *J. Phys. Chem.* **97,** 1101–1109
39. Keller, R. (2004) *The Computer Aided Resonance Assignment Tutorial, CANTINA*, Verlag, Goldau, Switzerland
40. Ottiger, M., Delaglio, F., and Bax, A. (1998) *J. Magn. Reson.* **131,** 373–378
41. Wang, Y. X., Marquardt, J. L., Wingfield, P., Stahl, S. J., Lee-Huang, S., Torchia, D., and Bax, A. (1998) *J. Am. Chem. Soc.* **120,** 7385–7386
42. Dosset, P., Hus, J. C., Marion, D., and Blackledge, M. (2001) *J. Biomol. NMR* **20,** 223–231
43. Zweckstetter, M., and Bax, A. (2000) *J. Am. Chem. Soc.* **122,** 3791–3792
44. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
45. Collaborative Computation Project (1994) *Acta Crystallogr. D. Biol. Crystallogr.* **50,** 760–763
46. Navaza, J. (1994) *Acta Crystallogr. A.* **50,** 157–163
47. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D. Biol. Crystallogr.* **53,** 240–255
48. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D. Biol. Crystallogr.* **60,** 2126–2132
49. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D. Biol. Crystallogr.* **54,** 905–921
50. Linge, J. P., and Nilges, M. (1999) *J. Biomol. NMR* **13,** 51–59
51. Garcia, M. M., Jimenez Rios, M. A., and Garcia Bernal, J. M. (1990) *Int. J. Biol. Macromol.* **12,** 19–24

52. Lo Conte, L., Chothia, C., and Janin, J. (1999) *J. Mol. Biol.* **285,** 2177–2198
53. Ramanathan, R. K., Trump, D. L., Eiseman, J. L., Belani, C. P., Agarwala, S. S., Zuhowski, E. G., Lan, J., Potter, D. M., Ivy, S. P., Ramalingam, S., Brufsky, A. M., Wong, M. K. K., Tutchko, S., and Egorin, M. J. (2005) *Clin. Cancer Res.* **11,** 3385–3391
54. Gray, P. J., Jr., Prince, T., Cheng, J., Stevenson, M. A., and Calderwood, S. K. (2008) *Nat. Rev. Cancer* **8,** 491–495
55. Zhang, T., Hamza, A., Cao, X., Wang, B., Yu, S., Zhan, C. G., and Sun, D. (2008) *Mol. Cancer Ther.* **7,** 162–170
56. Kaelin, W. G., Jr. (2005) *Nat. Rev. Cancer* **5,** 689–698
57. Takahashi, H., Nakanishi, T., Kami, K., Arata, Y., and Shimada, I. (2000) *Nat. Struct. Biol.* **7,** 220–223

8 Small Molecule Inhibitor for the Protein-Protein (Cdc37-Hsp90) Complex Studied by NMR

8.1 Research Article

Molecular Mechanism of Inhibition of the Human Protein Complex Hsp90-Cdc37, A Kinome Chaperone-Cochaperone, by Triterpene Celastrol

Sridhar Sreeramulu, Santosh Lakshmi Gande, Michael Göbel, Harald Schwalbe, *Angewandte Chemie International Edition*, 2009, *In press*.

In this article we described about a small molecule inhibitor for disrupting the protein-protein (Hsp90-Cdc37) complex.

Here we show by NMR how celastrol, a recently identified triterpene targeting Hsp90, in fact binds to Cdc37 and disrupts this tightly binding (k_D = 1.2 μM) complex. Further we also delineate as to which domains in Cdc37 are the targets for the binding of celastrol and also give experimental evidence for the exact mechanism of action.

The author of this thesis has performed all the UV, CD, and NMR experiments and analysis. The molecular biology work was performed together with Dr. Santosh Lakshmi Gande.

Protein-protein complex inhibitor

DOI: 10.1002/anie.200((will be filled in by the editorial staff))

Molecular Mechanism of Inhibition of the Human Protein Complex Hsp90-Cdc37, a Kinome Chaperone-Cochaperone, by Triterpene Celastrol**

*Sridhar Sreeramulu, Santosh Lakshmi Gande, Michael Göbel and Harald Schwalbe**

Inhibition of the ATPase activity of the kinome chaperone Hsp90 has long been known as molecular target for anticancer therapy. Cdc37, a cochaperone of Hsp90 in mammalian cells, targets protein kinases and is upregulated in various cancers.[1, 2] The protein-protein complex forms with a K_D = 1.2 μM, and is considered to mediate carcinogenesis by stabilizing a variety of different oncogenic kinases in malignant cells. However, Cdc37 as well as Hsp90 can act also alone, at least in yeast. Low molecular weight molecules that interfere with Cdc37 or Hsp90 or disrupt the Hsp90-Cdc37 complex have recently been proposed as a new class of anti-cancer agents.[3, 4] Gene based-expression studies have identified the triterpene celastrol, representing a new class of non ATP-competitive inhibitors of Hsp90.[5] Immunoprecipitation in a pancreatic cell line and docking experiments suggested that celastrol exerts its anti-proliferative activity by binding to the N-terminal domain of Hsp90 ($Hsp90_N$), thereby disrupting the complex between $Hsp90_N$ and Cdc37.[6] *In vivo*, celastrol showed significant inhibition of tumor growth in nude mice with prostate or pancreatic cancer.[6, 7]

In this communication, we investigate in detail how celastrol disrupts the human Hsp90-Cdc37 complex. $^1H,^{15}N$-HSQC (heteronuclear spin quantum correlation) NMR (nuclear magnetic resonance) experiments detect ligand binding to a target through chemical-shift perturbations (CSPs).[8-10] We investigated the effects of celastrol on the complex of $^1H,^{15}N$-labeled $Hsp90_N$ (23 kDa) with unlabeled full-length Cdc37 (45 kDa) (Figure 1A). Due to the increased transverse relaxation rates of the protein in complex (ca. 70 kDa), the HSQC spectrum of the complex resulted in broadening of most of the cross peaks, indicating complex formation. Addition of celastrol dissociated the complex and restored the HSQC spectrum of free $Hsp90_N$. Surprisingly, the HSQC spectrum of free $Hsp90_N$ and the spectrum obtained after the addition of celastrol to the $Hsp90_N$-Cdc37 complex matched exactly, revealing that celastrol disrupted the complex by a mechanism other than binding to $Hsp90_N$. Changes in chemical shifts indicative of binding upon the addition of celastrol to free $Hsp90_N$ could not be observed (see Supporting Information 1 (SI 1)).

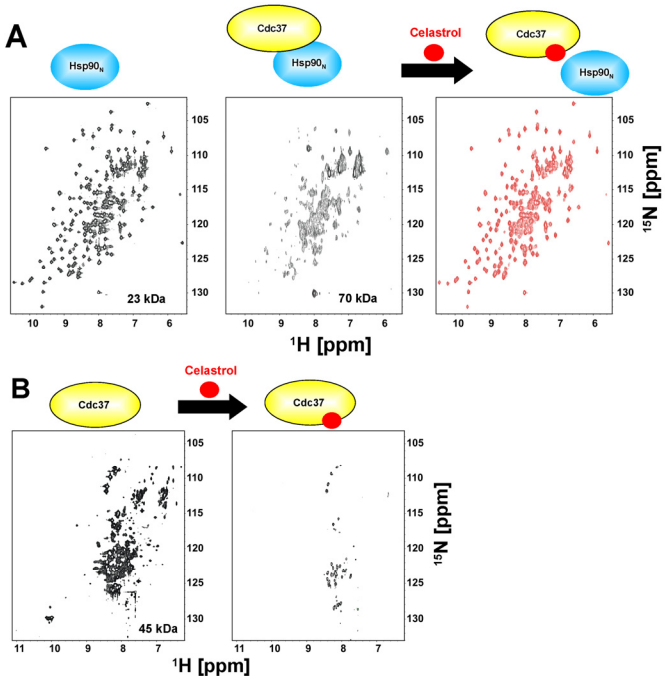


Figure 1. Binding of celastrol to Cdc37. A) (left) $^1H,^{15}N$ HSQC spectrum of $Hsp90_N$. (middle) $^1H,^{15}N$ HSQC spectrum of $Hsp90_N$-Cdc37 complex. (right) Upon addition of celastrol (1:7 of $Hsp90_N$-Cdc37:celastrol), the complex dissociates and the spectrum of free $Hsp90_N$ is retained indicating that celastrol does not bind to $Hsp90_N$. B) (left) $^1H,^{15}N$ HSQC spectrum of full-length Cdc37 (45 kDa). (right) The cross-peaks disappear upon addition of celastrol (1:7 of Cdc37:celastrol), indicating a complex formation with Cdc37.

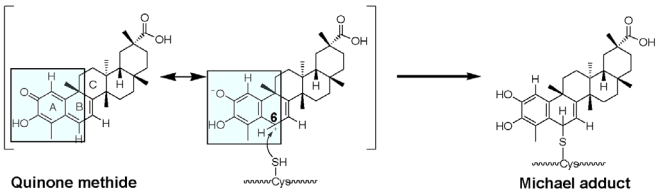
Then, we studied the interaction between Cdc37 and celastrol. For this purpose, we recorded a $^1H,^{15}N$-HSQC spectrum of ^{15}N-labeled Cdc37 (Figure 1B). Upon addition of celastrol to Cdc37, most of the cross peaks in the HSQC disappeared indicating interaction of the protein with the ligand. The loss of resonance signals in the HSQC spectrum could be due to aggregation of protein induced by the ligand. Celastrol has an orange color and becomes colorless upon addition to Cdc37, which suggests a loss of its chromophore as quantified by UV spectroscopy [11] (see SI 2).

Celastrol possesses electrophilic sites within the A and B ring (Scheme 1), where nucleophilic groups of amino acid residues react to form covalent Michael adducts.[7, 12, 13] In fact, mass spectrometry of Cdc37 in the presence of celastrol exhibited a mass increase of 450 Da, which corresponds to the molecular weight of celastrol (see SI 3).

[*] M. Pharm. S. Sreeramulu, Dr. S. L. Gande, Prof. Dr. M. Göbel, Prof. Dr. H. Schwalbe
Johann Wolfgang Goethe-University, Frankfurt am Main, Institute for Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance (BMRZ), Max-von-Laue-Strasse 7, D-60438 Frankfurt am Main, Germany
Fax: (+)49 69 79829515
E-mail: schwalbe@nmr.uni-frankfurt.de
Homepage: http://schwalbe.org.chemie.uni-frankfurt.de/

[**] Cdc37: Cell division cycle protein 37, Hsp90: Heat shock protein 90. We thank Mr. Frédéric tournay at the Jardin Botanique de l'Universite Louis Pasteur, France, for kindly providing the photograph of the plant *Tripterygium wilfordii*. This work was supported by the SPINEII project of the European commission and by the Cluster of Excellence: Macromolecular Complex (DFG)

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.



Scheme 1. Molecular mechanism for the addition of celastrol to thiol of Cdc37. The quinone methide region of celastrol is highlighted in a blue box in the structure. The thiol of Cdc37 reacts with the α,β-unsaturated ketone of celastrol through Michael addition reaction, resulting in the formation of an adduct at C6.

To confirm that celastrol modifies the cysteine residue(s) in Cdc37, we blocked all nine cysteine residues of full-length Cdc37 with N-ethylmaleimide (NEM). NEM-labeled Cdc37 no longer reacted with celastrol and the mixture did not show any change in the UV spectrum indicating that cysteine(s) undergo(es) chemical reactions with celastrol upon binding (see SI 4). Circular dichroism (CD) spectra for the wild-type Cdc37 and Cdc37_NEM showed a decrease in secondary structure, indicating that covalent modification of the free cysteines alters the conformation of the protein (see SI 5). Further, the HSQC spectrum of $Hsp90_N$ remained unchanged upon addition of Cdc37_NEM, indicating that modification of thiols is sufficient to inhibit the interaction between $Hsp90_N$ and Cdc37. Free cysteine can undergo a covalent addition reaction with celastrol, as the lipophilic celastrol becomes water soluble after addition of cysteine at a pH of 7.4, loses its absorption at 440 nm and reveals chemical shift changes in 1D proton NMR, characteristic for the Michael adduct at C6 in ring B of the triterpene (see SI 6).

Michael adduct formation using cysteine thiol groups as nucleophile is highly dependent on pH. Under physiological conditions, the reaction is fast and reversible, thereby the ligand escapes from high GSH-concentration present in living cells (0.5-10 mM) and reaches the target protein.[14]

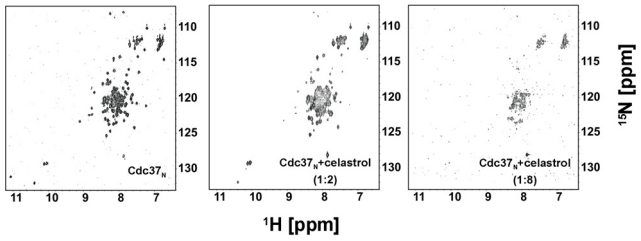


Figure 2. Binding of celastrol to $Cdc37_N$. (left) ^1H,^{15}N HSQC spectrum of free $Cdc37_N$. (middle) Some of the cross-peaks disappear upon addition of celastrol that forms a complex with $Cdc37_N$ (1:2). (right) Most of the cross-peaks disappear upon further addition of celastrol to $Cdc37_N$ (1:8).

Cdc37 is composed of three domains, the N-terminal kinase binding domain ($Cdc37_N$, residues 1-126) with three free cysteines, the middle, $Hsp90_N$ binding domain ($Cdc37_M$, residues 147-276) with four free cysteines and the C-terminal domain ($Cdc37_C$, residues 282-376) of unknown function with two free cysteines (see SI 7). In order to delineate which of these domains is responsible for binding celastrol, individual domain constructs were expressed. $Cdc37_N$ reduced celastrol and resulted in an inactivated form, as observed in the ^1H,^{15}N-HSQC spectrum (Figure 2). $Cdc37_M$ alone precipitated upon addition of celastrol, but did not reduce the compound, while the construct with the middle and the C-terminal domain ($Cdc37_{MC}$) reacted with celastrol, but also precipitated. This observation suggests that the N-terminal domain is mainly responsible for reducing celastrol and binding. Additionally, when the cysteines were blocked using NEM, the $Cdc37_N$ no longer interacted with celastrol (see SI 8).

Further, in order to see which of the three cysteines of $Cdc37_N$ are involved in the interaction with celastrol, we made double cysteine mutants retaining one cysteine ($Cdc37_N$-C57S-C64S, $Cdc37_N$-C54S-C64S, $Cdc37_N$-C54S-C57S). Circular dichroism (CD) spectra for the wild-type and mutant $Cdc37_N$ are mostly similar, indicating that there are no changes in secondary structure (see SI 5). We monitored the decrease in absorbance at 440 nm after the addition of various constructs of Cdc37 to celastrol. $Cdc37_N$ and all the three mutants reduced celastrol and were comparable to the wild-type (see SI 4), indicating they are all reactive.

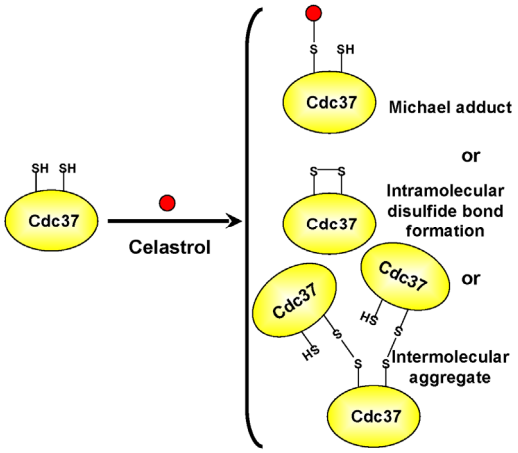


Figure 3. Mechanism of inactivation of Cdc37 by celastrol. Celastrol inactivates Cdc37 either by the formation of a covalent adduct(s) with thiol(s) of Cdc37, by formation of an intramolecular disulfide bond or by intermolecular aggregates.

In conclusion, we report that celastrol binds to Cdc37 but not to $Hsp90_N$ as previously described. This study is important as it will shift medicinal chemistry efforts towards the correct target within the protein-protein complex. Celastrol inactivates Cdc37 by covalently binding to it or by forming either an intra- or intermolecular protein disulfide (Figure 3). Reactivities of individual cysteines varies slightly (as seen from the UV results), depending on accessibility and the stabilization interaction of the adjacent amino acids, celastrol binds to one of the cysteines. The interaction between Cdc37-Hsp90 also involves a large hydrophobic core at the interface.[15] Adjacent to the hydrophobic core is C203 from $Cdc37_M$, which is exposed to the surface. Celastrol harbors a hydrophobic structure apart from the quinone methide moiety. Hence, celastrol could be involved disrupting this hydrophobic core. The binding of celastrol induces large changes in conformation of the N-terminal kinase binding domain and also the middle $Hsp90_N$ binding domain of Cdc37, thereby disrupting the Cdc37-$Hsp90_N$ complex which is crucial for stabilizing oncogenic kinases in various cancers.[3] Our results indicates that the N-terminal kinase binding and middle $Hsp90_N$ domain of Cdc37 are the molecular target for the triterpene

celastrol. In light of the considerable potential of Cdc37 as a drug target and binding of celastrol to it, these results can now be used to further understand the molecular basis for inhibiting this key Cdc37-Hsp90 complex in cancer cells.

Experimental Section

(see Supporting Information)

Received: ((will be filled in by the editorial staff))
Published online on ((will be filled in by the editorial staff))

Keywords: inhibitors · NMR spectroscopy · chaperone · Hsp90 · Cdc37 ·

[1] P. Lee, J. Rao, A. Fliss, E. Yang, S. Garrett, A. J. Caplan, *J. Cell Biol.* **2002**, *159*, 1051.
[2] L. H. Pearl, *Curr. Opin. Genet. Dev.* **2005**, *15*, 55.
[3] P. J. Gray, Jr., T. Prince, J. Cheng, M. A. Stevenson, S. K. Calderwood, *Nat. Rev. Cancer* **2008**, *8*, 491.
[4] J. R. Smith, P. A. Clarke, E. de Billy, P. Workman, *Oncogene* **2009**, *28*, 157.
[5] H. Hieronymus, J. Lamb, K. N. Ross, X. P. Peng, C. Clement, A. Rodina, M. Nieto, J. Du, K. Stegmaier, S. M. Raj, K. N. Maloney, J. Clardy, W. C. Hahn, G. Chiosis, T. R. Golub, *Cancer Cell* **2006**, *10*, 321.
[6] T. Zhang, A. Hamza, X. Cao, B. Wang, S. Yu, C. G. Zhan, D. Sun, *Mol. Cancer Ther.* **2008**, *7*, 162.
[7] H. Yang, D. Chen, C. C. Qiuzhi, X. Yuan, Q. P. Dou, *Cancer Res.* **2006**, *66*, 4758.
[8] L. D'Silva, P. Ozdowy, M. Krajewski, U. Rothweiler, M. Singh, T. A. Holak, *J. Am. Chem. Soc.* **2005**, *127*, 13220.
[9] C. A. Lepre, J. M. Moore, J. W. Peng, *Chem. Rev.* **2004**, *104*, 3641.
[10] B. J. Stockman, C. Dalvit, *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *41*, 187.
[11] K. T. Liby, M. M. Yore, M. B. Sporn, *Nat. Rev. Cancer* **2007**, *7*, 357.
[12] A. Trott, J. D. West, L. Klaic, S. D. Westerheide, R. B. Silverman, R. I. Morimoto, K. A. Morano, *Mol. Biol. Cell* **2008**, *19*, 1104.
[13] J. H. Lee, T. H. Koo, H. Yoon, H. S. Jung, H. Z. Jin, K. Lee, Y. S. Hong, J. J. Lee, *Biochem. Pharmacol.* **2006**, *72*, 1311.
[14] T. J. Schmidt, G. Lyss, H. L. Pahl, I. Merfort, *Bioorg. Med. Chem.* **1999**, *7*, 2849.
[15] S. Sreeramulu, H. R. Jonker, T. Langer, C. Richter, C. R. Lancaster, H. Schwalbe, *J. Biol. Chem.* **2009**, *284*, 3885.

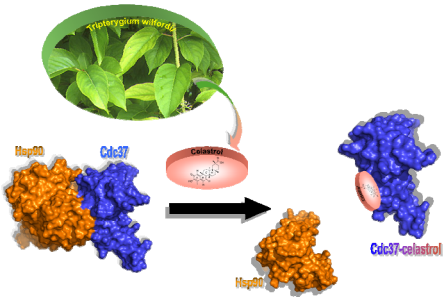
Entry for the Table of Contents

Protein-protein complex inhibitor

*S. Sreeramulu, S. L. Gande, M. Göbel, H. Schwalbe **

Page – Page

Molecular Mechanism of Inhibition of the Human Protein Complex Hsp90-Cdc37, a Kinome Chaperone-Cochaperone, by Triterpene Celastrol



The correct target: The cell division cycle protein 37 (Cdc37) and the heat shock protein (Hsp90) are molecular chaperones, crucial for the folding and stabilization of protein kinases including the oncogenic kinases. Here we show by NMR that celastrol, a recently identified triterpene targeting Hsp90, in fact binds to Cdc37 and disrupts the Cdc37-Hsp90 complex. Celastrol inactivates Cdc37 through a thiol-mediated mechanism.

Supporting Information

Protein-protein complex inhibitor

Molecular Mechanism of Inhibition of the Human Protein Complex Hsp90-Cdc37, a Kinome Chaperone-Cochaperone, by Triterpene Celastrol

Sridhar Sreeramulu, Santosh Lakshmi Gande, Michael Göbel and Harald Schwalbe

Johann Wolfgang Goethe-University, Frankfurt am Main, Institute for Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance (BMRZ), Max-von-Laue-Strasse 7, D-60438 Frankfurt am Main, Germany
Fax: (+)49 69 79829515
E-mail: schwalbe@nmr.uni-frankfurt.de
Homepage: http://schwalbe.org.chemie.uni-frankfurt.de/

Experimental Section

Protein expression and purification: The human Cdc37, $Cdc37_N$ and $Hsp90_N$ was expressed in *E. coli* strain BL21 (DE3) and purified by nickel affinity chromatography as described previously.[15] The uniformly ^{15}N enriched protein samples were prepared by growing the bacteria in minimal media containing ^{15}N-NH_4Cl. After the purification the protein was dialyzed against a buffer containing 50 mM Hepes pH 7.4, 100 mM NaCl, 50 µM of TCEP, concentrated (approximately to 1 mM) and stored at -20 °C.
Mutagenesis: Mutants of $Cdc37_N$ were created with QuickChange kit (Stratagene), and sequenced to confirm the incorporation of the correct mutation. The mutant proteins were purified following the same protocol as the wild-type.
NMR spectroscopy: NMR experiments were performed at 298 K, on Bruker 600, 800, 900 and 950 MHz spectrometers equipped with cryogenic triple-resonance probes. For the Cdc37-$Hsp90_N$ complex a ratio of 1:1 was used for all measurements. Typically, NMR samples contained 0.6 ml of up to 0.15 mM protein in 50 mM Hepes pH 7.4, 100 mM NaCl, 0.05 % NaN_3. The spectra were processed using Topspin 2.0 (Bruker Biospin) and analyzed using SPARKY 3.113.
NMR titrations with celastrol: Celastrol was purchased from Shangai Yingxuan Pharmaceutical Science & Technology Co., LTD. Since celastrol is not soluble in the buffer used for NMR samples (as judged by the 1D ^1H NMR spectrum), DMSO-d_6 stock solution of celastrol was used for NMR measurements. The effect of DMSO-d_6 on the proteins and the protein complex was checked to be negligible up to 3 % v/v of DMSO-d_6.
To check whether celastrol might bind to $Hsp90_N$, we carried out titration of celastrol with ^{15}N $Hsp90_N$ and could positively determine that celastrol does not bind to $Hsp90_N$. No major changes in the $Hsp90_N$ spectrum were observed even up to a ratio of 1:7 ($Hsp90_N$: celastrol). In similar fashion experiments were carried out for Cdc37, $Cdc37_N$ and the $Hsp90_N$-Cdc37 complex.
For studying the interaction between free cysteine and celastrol, we performed the reaction in Phosphate buffer saline pH 7.4. Celastrol and free cysteine was mixed in a ratio of 1:20 and then recorded the 1D ^1H NMR spectrum. Further, on long standing of the mixture of cysteine (or N-acetyl cysteine or GSH) and celastrol, one observes a white precipitate due to the polymerization of the reduced celastrol.
N-Ethylmaleimide (NEM) labelling: Free cysteine residues in Cdc37/$Cdc37_N$ were blocked using NEM reagent (Pierce Biotechnology, Inc., Rockford, IL). The reaction was carried out in a buffer containing 50 mM Hepes pH 7.2, 100 mM NaCl. Cdc37 was allowed to react with NEM (1:10) for 2 hrs at room temperature, subsequently, the excess NEM was removed by dialysis (Slide-A-Lyzer). The NEM labeling was confirmed by MALDI-TOF mass spectrometry.

MALDI-TOF mass spectrometry analysis: Cdc37 (in our case we used the ^{15}N labeled protein) was incubated with celastrol (1:7) in a buffer containing 50 mM Hepes pH 7.4, 100 mM NaCl, 0.05 % NaN_3 for 10 minutes. A MALDI-TOF mass spectrometry spectrum was then obtained with a Voyager Elite.

Measurement of UV spectra: Samples were mixed at indicated concentrations in a buffer containing 50 mM Hepes pH 7.4, 100 mM NaCl and then monitored between 200 and 600 nm by using UV spectrophotometer (Cary 50 Bio UV-visible spectrophotometer, Varian). All spectra were later smoothened (Savitzky-Golay method) and plotted using Origin (version 8).

Measurement of circular dichroism (CD) spectra: All measurements were carried out with a Jasco J-810 CD spectropolarimeter in a cuvette with a path length of 0.1 cm. The scanning speed was 50 nm/min, bandwidth 0.1 nm, and a response time of 1 s. In each experiment, measurements were done at 25 °C and 3 spectra were summed and averaged. All protein samples were typically either 5 μM or 15 μM in a buffer containing 5 mM Hepes pH 7.4, 10 mM NaCl. The concentration of protein was measured by using Nanodrop spectrophotometer, ND-1000, PEQLAB.

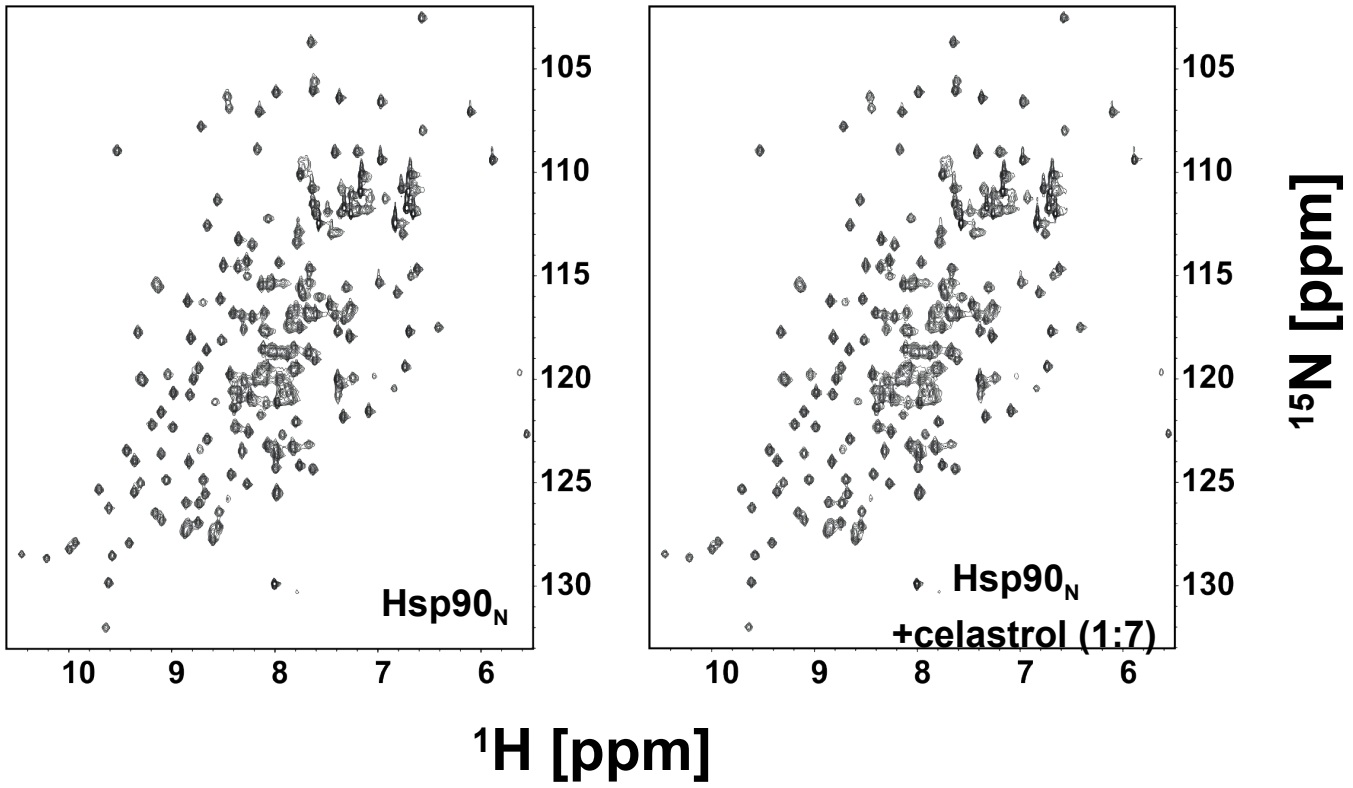


Figure S1. Celastrol does not bind to $Hsp90_N$. (left) $^{1}H,^{15}N$ HSQC spectrum of free $Hsp90_N$. (right) $^{1}H,^{15}N$ HSQC spectrum of $Hsp90_N$ plus celastrol (1:7). Both spectra remain identical, indicating that celastrol does not bind to $Hsp90_N$.

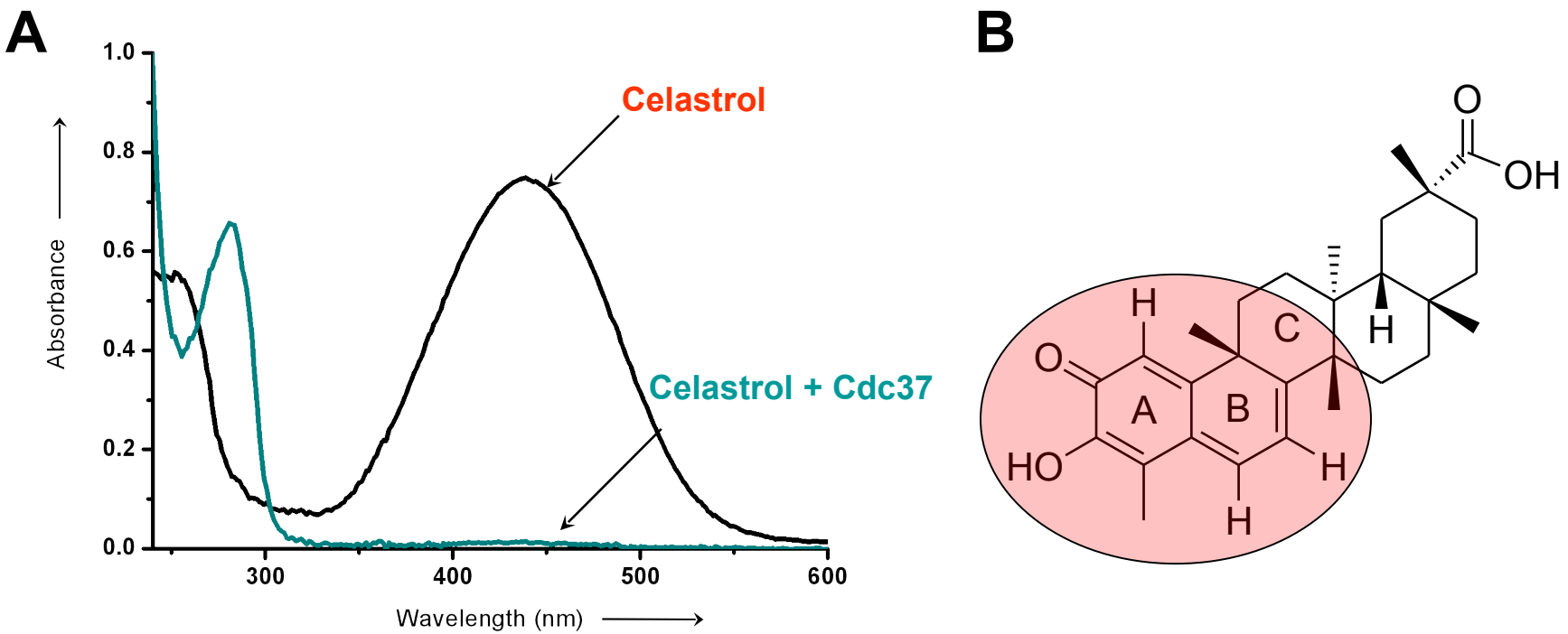


Figure S2. Change of UV spectrum of celastrol upon binding to Cdc37. A) UV spectrum of 100 μM of celastrol before (black) and after incubation with 10 μM Cdc37 for 30 minutes (cyan). B) Structure of celastrol with the chromophore region shown in pink.

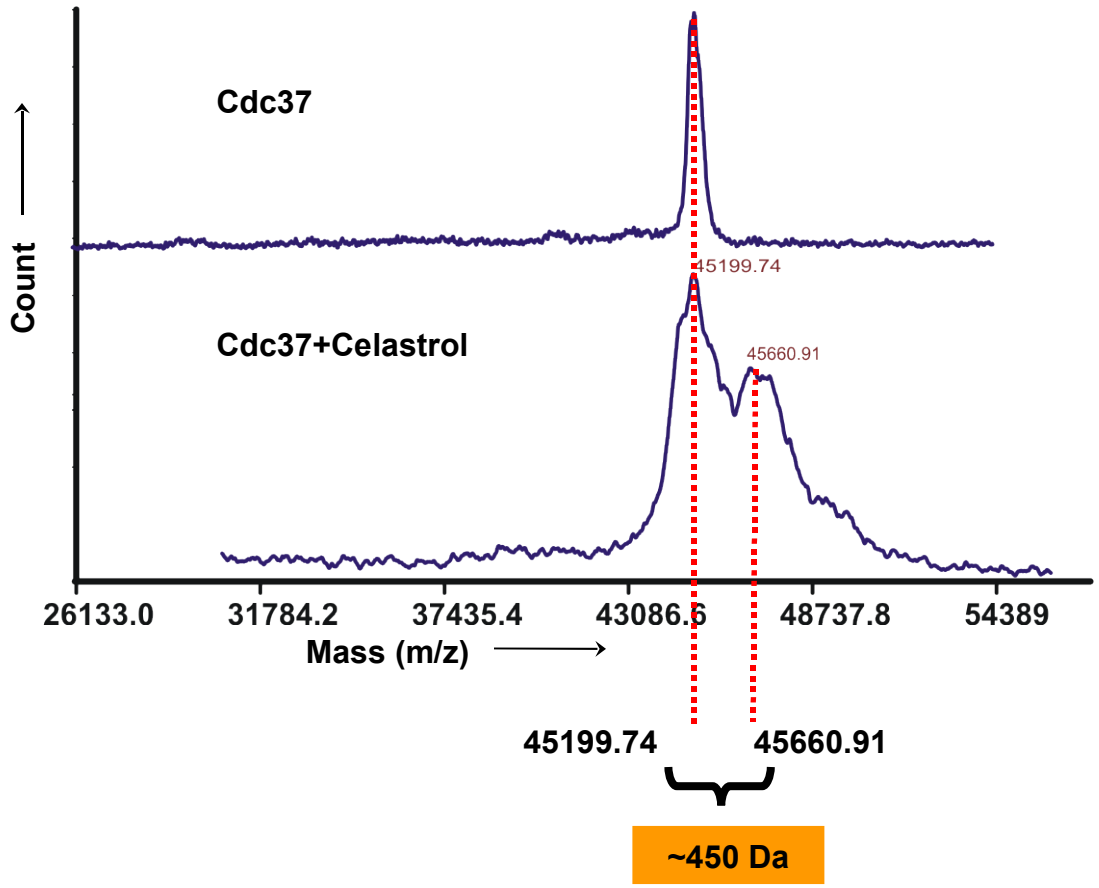
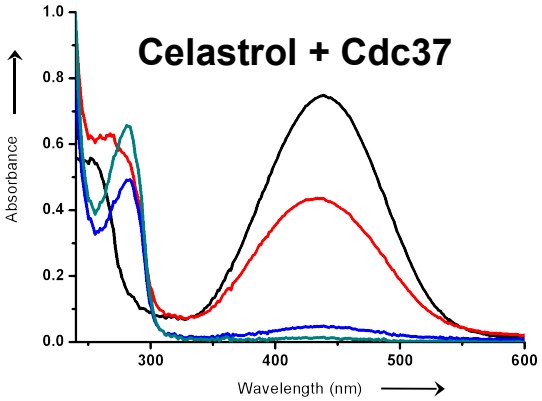
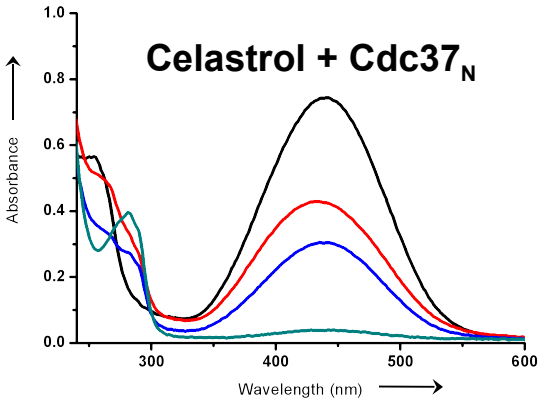
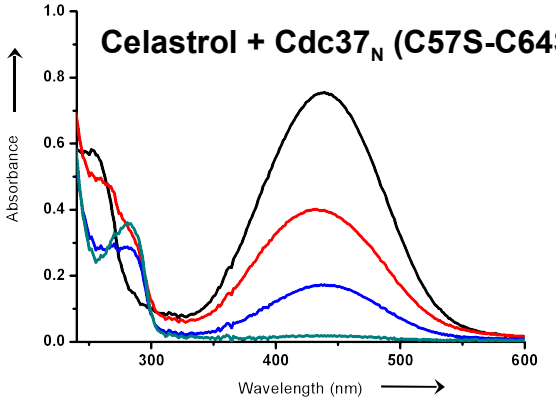
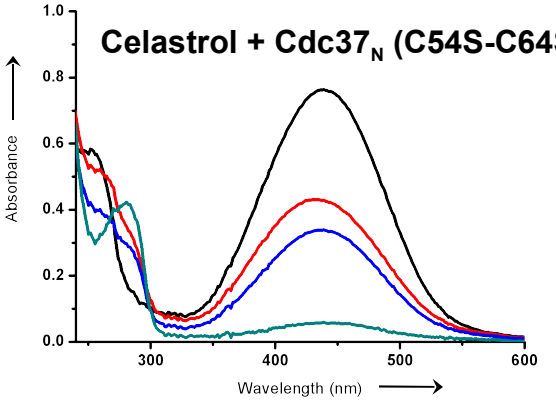


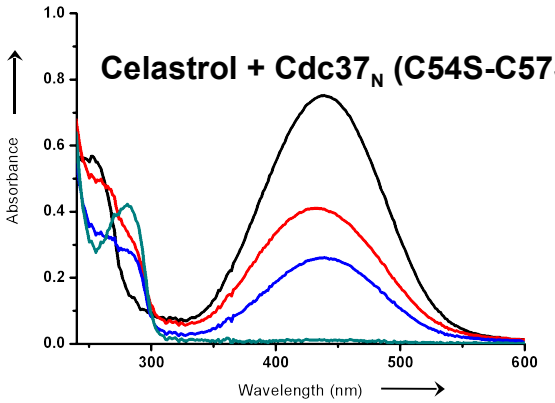
Figure S3. Covalent binding of celastrol to Cdc37. Mass spectrometry analysis of Cdc37 with and without celastrol. After incubating Cdc37 with celastrol, the molecular weight of the Cdc37-celastrol complex was investigated by MALDI-TOF. The molecular mass difference of 450 Da suggests that celastrol (Molecular weight: 450 Da) is covalently bound to Cdc37.

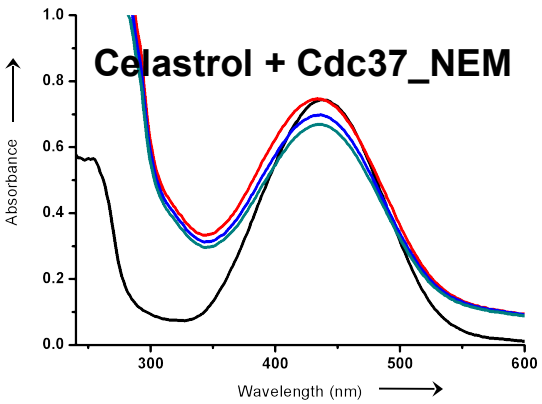

Celastrol + Cdc37
Absorbance
Wavelength (nm)

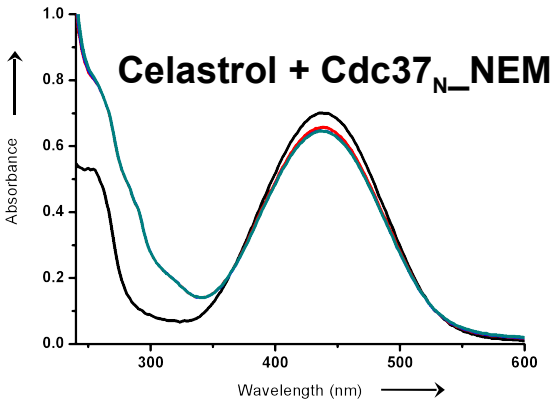

Celastrol + Cdc37N
Absorbance
Wavelength (nm)


Celastrol + Cdc37N (C57S-C64S)
Absorbance
Wavelength (nm)


Celastrol + Cdc37N (C54S-C64S)
Absorbance
Wavelength (nm)


Celastrol + Cdc37N (C54S-C57S)
Absorbance
Wavelength (nm)


Celastrol + Cdc37_NEM
Absorbance
Wavelength (nm)


Celastrol + Cdc37N_NEM
Absorbance
Wavelength (nm)

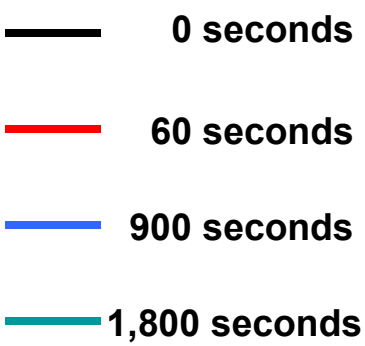
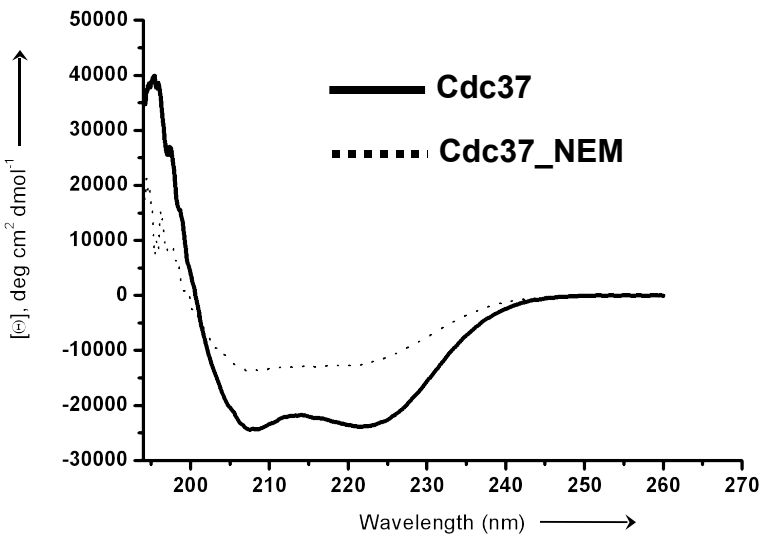

0 seconds
60 seconds
900 seconds
1,800 seconds

Figure S4. UV absorption spectra of celastrol monitored after addition of Cdc37, $Cdc37_N$ and the three cysteine mutants of $Cdc37_N$. The decrease in absorbance of celastrol (100 μM, black) at 440 nm after the addition of various constructs of Cdc37 was monitored by UV (7 μM of Cdc37, 7 μM of $Cdc37_N$ and mutants). As seen from the graph, the reaction proceeds in a similar fashion for all the constructs during the first 60 seconds (red) and varies thereafter, indicating that the cysteines of $Cdc37_N$ are as reactive as in the Cdc37 full length protein. After 900 seconds the absorbance becomes stable, and the addition of another 3 μM (Cdc37) or 7 μM ($Cdc37_N$ or mutants) of protein results in complete reduction of celastrol (cyan). If we block the cysteines with NEM (last two spectra), Cdc37 is no longer able to react with celastrol, endorsing the fact that free cysteines are the clear targets.



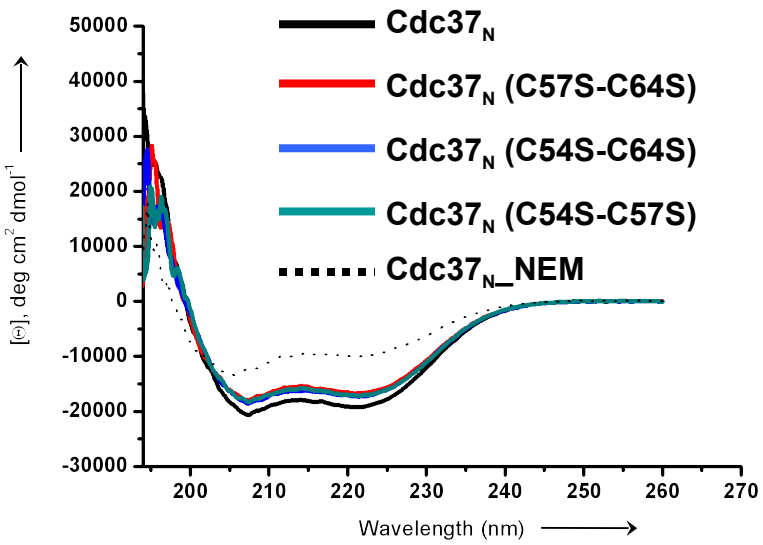


Figure S5. CD spectra of Cdc37, Cdc37$_N$ and Cdc37$_N$ mutants. (left) CD spectra of wild type Cdc37 (5 μM) and Cdc37 in which all the free cysteines are coupled to NEM. (right) CD spectra of Cdc37$_N$ (15 μM), Cdc37$_N$_NEM and the cysteine mutants. The spectra of the mutants show no significant changes compared to the wild type, indicating no change in secondary structure due to mutation.

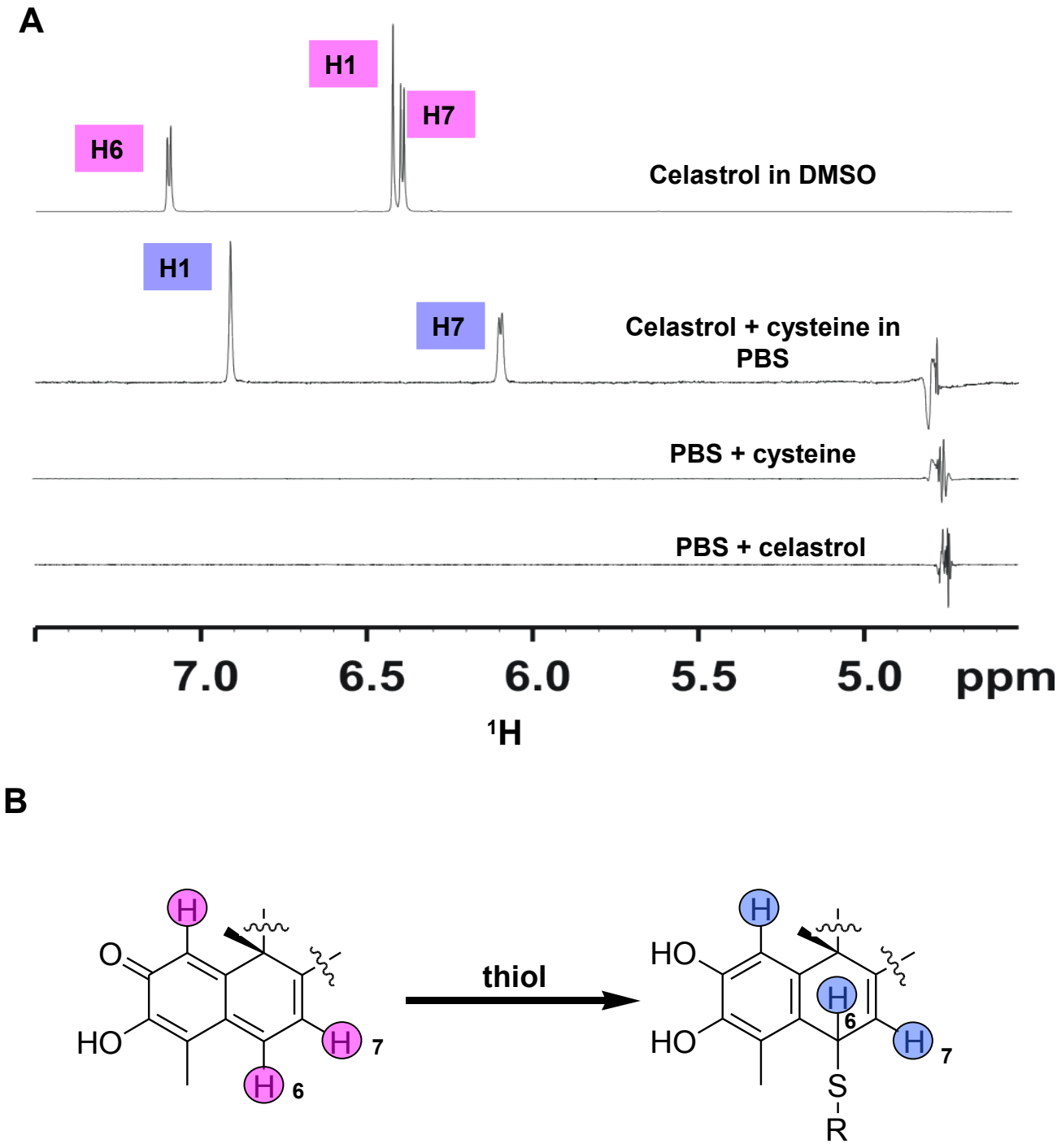


Figure S6. Celastrol forms Michael adduct with cysteine. A) Overlay of 1D ^{1}H NMR spectra of celastrol with and without cysteine in phosphate buffer pH 7.4 (PBS) / DMSO. Since celastrol is sparingly soluble in water, we do not see any signal for celastrol in PBS. In the 1D ^{1}H NMR spectrum of celastrol upon addition of cysteine, the insoluble celastrol now is reduced by cysteine and becomes water soluble. In the 1D ^{1}H NMR spectrum of celastrol with cysteine in PBS, the signals for the H1 and H7 protons appear after the reaction with cysteine. H6 in reduced celastrol is shifted and could not be detected because of its overlap with the water signal. As celastrol is soluble in DMSO, all signals (H1, H6 and H7) could be observed in the 1D ^{1}H NMR spectrum. B) Schematic representation of the conceived reaction.

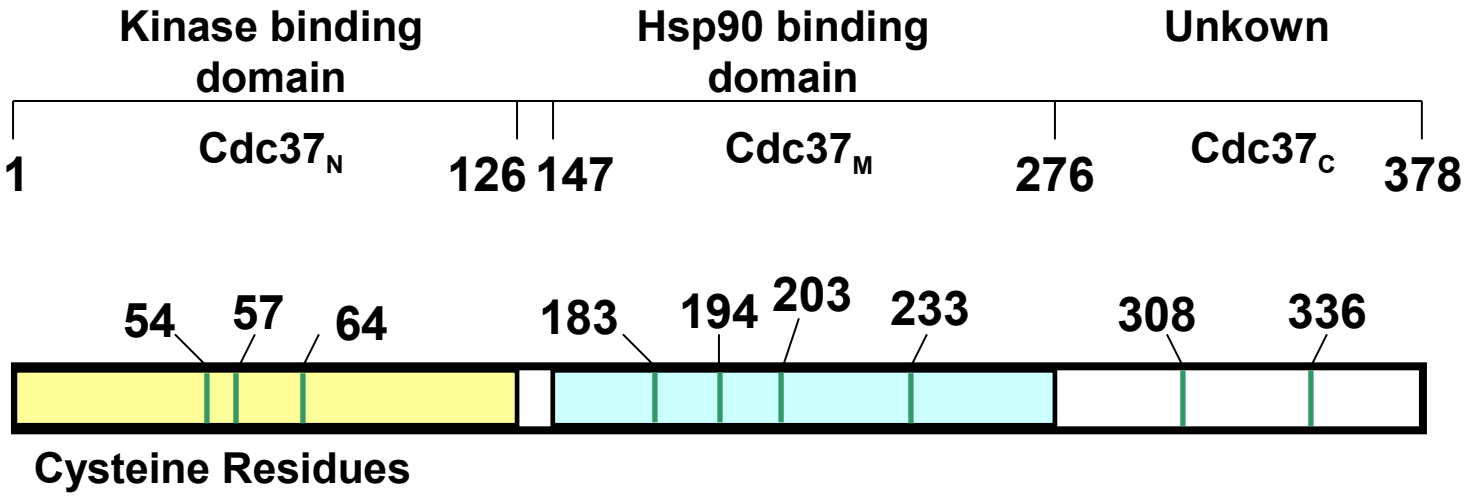


Figure S7. Schematic representation of the three domains of Cdc37 and the cysteine distribution.

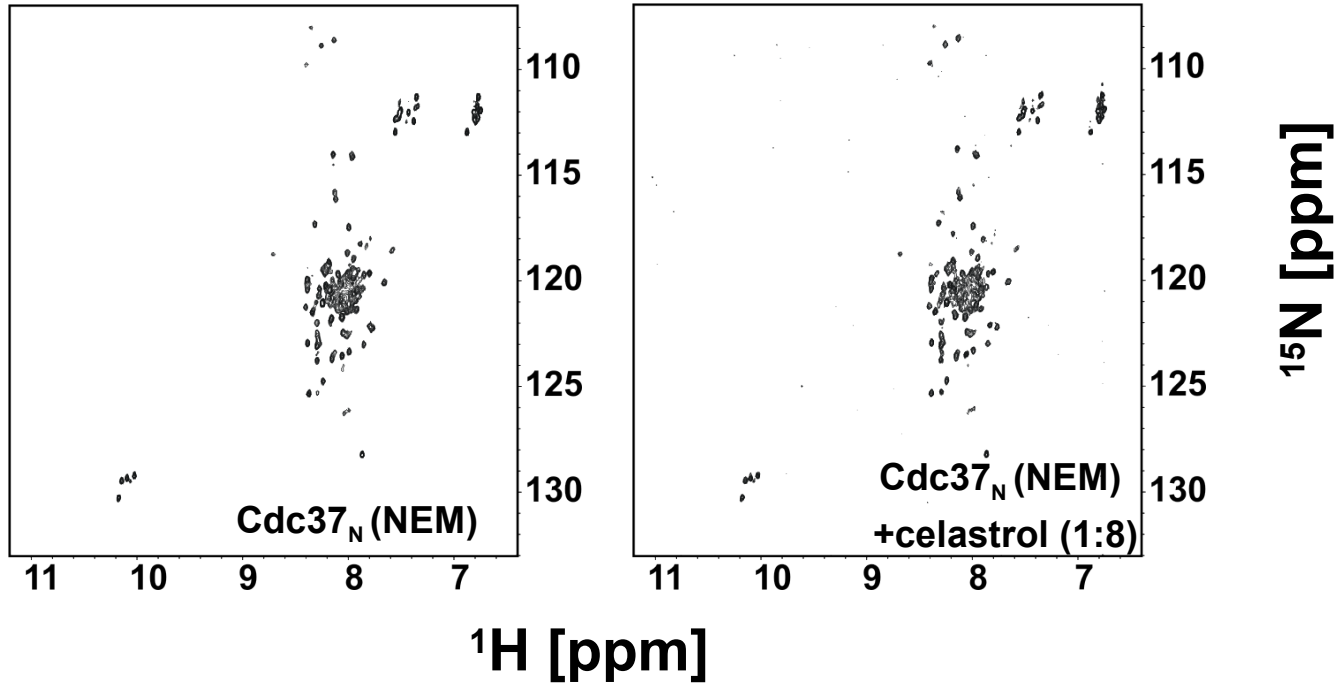


Figure S8. Cysteines in $Cdc37_N$ are responsible for the binding of celastrol. (left) $^{1}H,^{15}N$ HSQC spectrum of $Cdc37_N$ in which all three cysteines are covalently bound to NEM. (right) $^{1}H,^{15}N$ HSQC spectrum of $Cdc37_N$-NEM after the addition of celastrol (1:8). Both spectra remains the same, indicating that there is no complex formation with celastrol, endorsing the fact that cysteines are the main target of celastrol.

ZUSAMMENFASSUNG UND ÜBERBLICK

Die folgende Arbeit beschäftigt sich mit der Untersuchung von Proteinen und Protein-Protein-Komplexen von biomedizinischer Bedeutung mittels Kernmagnetresonanzspektroskopie (NMR-Spektroskopie). Zu verstehen, wie biologische Systeme arbeiten von der Stufe eines einzelnen Proteins hin zu komplexeren Protein-Protein-Komplexen und letztlich zur dadurch vermittelten Weitergabe eines Signals, ist ein wichtiger Schritt, um Heilmittel für viele Krankheiten zu finden. Krebs ist eine solche Krankheit, die durch unkontrollierte Zellteilung, die oft durch Proteinkinasen gefördert wird, charakterisiert ist [Teicher, 2000]. Daher werden Proteinkinasen oftmals als Rezeptoren der Bindung kleiner Inhibitormoleküle benutzt. Dabei wird oft das Problem mangelnder Zielselektivität beobachtet, verursacht durch die hohe Sequenzähnlichkeit zwischen unterschiedlichen Mitgliedern der Kinasen-Familie [Fabbro et al., 2002; Noble et al., 2004].

Akt/PKB ist eine Proteinkinase, die häufig in einer grundlegend aktiven Form in humanen Krebsarten gefunden wurde. Akt/PKB ist ein klinisch validiertes Ziel und daher ist die Entwicklung kleiner Inhibitormoleküle gegen diese Kinase höchst erstrebenswert [Brognard et al., 2001; Hanada et al., 2004].

Akt/PKB kann nicht einfach in *E. coli* exprimiert werden. Daher wurde die cAMP-abhängige Kinase (PKA), die eine hohe Sequenz-Homologie zu Akt/PKB aufweist, als so genannte Surrogatkinase für die Wirkstoffentwicklung benutzt.

Alternativ dazu befinden sich Wirkstoffe, die das molekulare Chaperon Hsp90

inhibieren, gerade in klinischen Studien. Ihr Wirkmechanismus beruht auf der Inhibition der Funktion des Hsp90, den Fehlfaltungs-induzierten Abbau vieler Kinasen fördern zu können [Whitesell and Lindquist, 2005]. Kürzlich wurde mit Cdc37 ein Co-Chaperon von Hsp90 in Säugetierzellen identifiziert, das an Proteinkinasen bindet. In verschiedenen Krebsarten ist die Expression von Cdc37 hochreguliert [Lee et al., 2002; Pearl, 2005]. Der Protein-Protein-Komplex (Hsp90-Cdc37) bildet sich mit einem K_D = 1,2 μM und wird durch die Stabilisierung einer Vielfalt von verschiedenen onkogenen Kinasen in bösartigen Zellen als krebsauslösend betrachtet.

Eine Kombination aus NMR-Spektroskopie und Röntgenkristallographie wurde in dieser Arbeit angewendet, um diese Proteine oder Protein-Protein-Komplexe in hoher struktureller Auflösung zu untersuchen. Diese Untersuchungen, die Wechselwirkungen in den Protein-Protein-Komplexen mit atomarer Auflösung zu bestimmen sind wichtig, da es uns so die Identifikation einerseits der ‚Hot Spots' der Protein-Protein-Wechselwirkungsgrenzfläche, sowie in der weiteren Entwicklung, der kleinen Inhibitor-Moleküle ermöglichen wird [Wells and McClendon, 2007].

Der erste Teil der Arbeit stellt dem Leser die Welt der strukturellen Proteomik und seine Bedeutung vor. Kapitel 2 diskutiert eine Vielzahl von Projekten, die in der Post-humanen-Genom-Ära in verschiedenen Ländern weltweit unternommen worden sind, um dreidimensionale Strukturen einer beachtlichen Gruppe von Proteinen eines Organismus erhalten, das ‚Proteom'. Kapitel 2 berichtet außerdem vom EU-geförderten Projekt ‚Structural Proteomics in Europe' (SPINE), das seine Untersuchungen auf Proteine und Protein-Protein-Komplexe von biomedizinischer Bedeutung begrenzt hat und zeigt explizit PDB-Statistiken von dreidimensionalen Strukturen mit großem Molekulargewicht, die durch NMR-Spektroskopie aufgeklärt wurden.

Die Bedeutung der Untersuchung von Proteinen, die einen Bezug zu Krebs haben, wie die Proteinkinasen Hsp90 und Cdc37, wird im Kapitel 3 vorgestellt. Kapitel 4 gibt dem Leser einen Einblick in die Bedeutung der Entwicklung kleiner Inhibitormoleküle zur Trennung von Protein-Protein-Interaktionen, um Krankheiten vorzubeugen. Kapitel 5 gibt letztlich einen Überblick über die verschiedenen NMR-Methoden, die für die Untersuchungen von Protein-Protein-Wechselwirkungen genutzt wurden. Dieses Kapitel führt den Leser außerdem in die Bedeutung von Protein-Protein-Interaktionen in zellulären Prozessen ein.

Der zweite Teil der Arbeit, der kumulativ ist, umfasst die durchgeführte Forschungsarbeit. Kapitel 6 beschreibt, wie NMR-Spektroskopie genutzt wurde, um Zugang zur korrekten Faltung der Proteinkinase A (PKA) zu erhalten. Weiterhin wird darüber berichtet, wie Mutationen der Phosphorylierungsorte in PKA die Expressionsebene, die Stabilität und die Aktivität der Kinase beeinflussen.

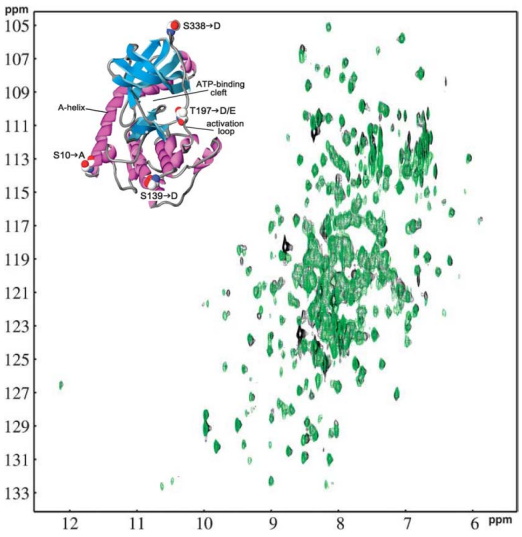


Abbildung 8.1: Faltung und Aktivität der cAMP-abhängigen Proteinkinasenmutanten. Der Vergleich der TROSY-Spektren des Wildtyps (schwarz) und der ‚dephopho'-PKA-Mutante, PKA-3P (grün), zeigt, dass die Mutanten korrekt gefaltet sind. Weitere Details sind in Kapitel 6 zu finden.

Ergebnisse:

- Etablierung von PKA mit drei Phospho-Mutationen, die in *E. coli* exprimiert werden kann, was sonst ein großes Hindernis zur Synthese der meisten humanen Kinasen ist
- Nachweis durch NMR-Spektroskopie, das die PKA mit den drei Phospho-Mutationen (40 kDa) korrekt gefaltet sind
- Nachweis, dass es keine Mischungen der Isoformen aufgrund unterschiedlicher Phosphorylierung gibt, da hier nur ein Phosphorylierungort im Akivierungs-Loop vorhanden ist. Dieses Proteinkonstrukt wird sehr hilfreich sein für Untersuchungen von Protein-Ligand-Interaktionen.
- Etablierung der Tatsache, dass diese Mutanten aktiv und korrekt gefaltet sind, was nun höchst vorteilhaft für weitere funktionelle und strukturelle Untersuchungen sein kann

Kapitel 7 erläutert, wie wir die Strukturvorhersage eines humanen Cdc37$_M$-Hsp90$_N$-Protein-Protein-Komplexes (40 kDa) mittels NMR-Spektroskopie in Kombination mit Röntgenkristallographie erzielen konnten. Weiterhin konnten

wir die Lage des ‚Hot Spots' in der großen Wechselwirkungsgrenzfläche des Protein-Protein-Komplexes genau feststellen.

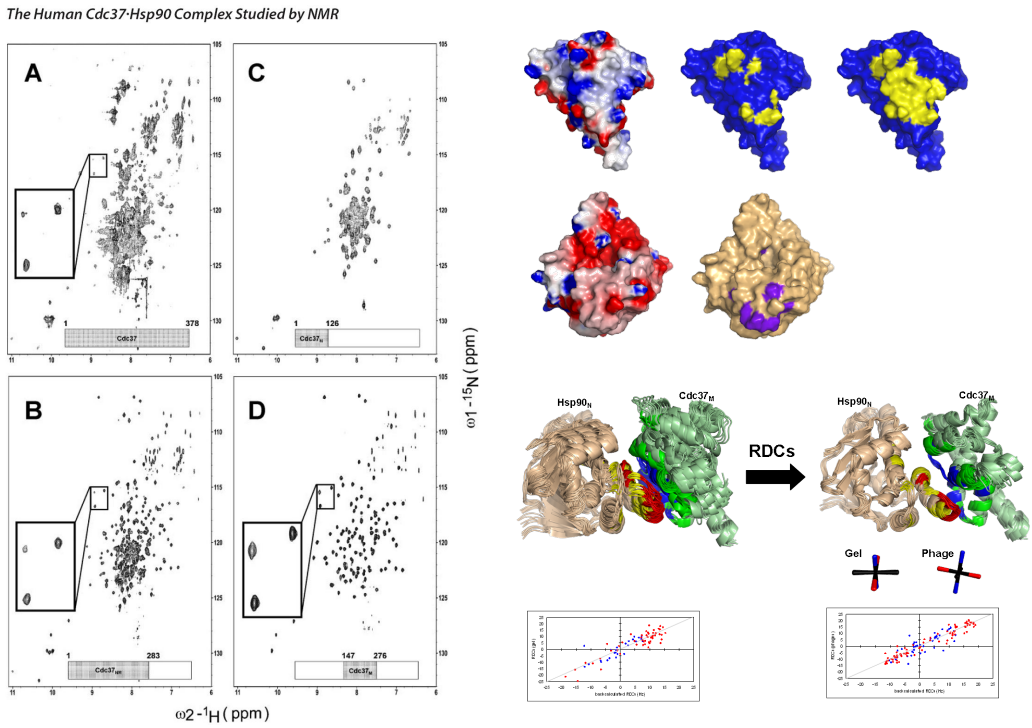


Abbildung 8.2: NMR-Studien des humanen Cdc37$_M$-Hsp90$_N$-Komplexes. Eine Kombination aus Domänenoptimierung, NMR-spektroskopischen Methoden (Resonanzzuordnung, Abbilden der chemischen Verschiebungsstörung, Kreuzsättigungstransfer, RDC's), Röntgenkristallographie und Docking mit HADDOCK wurde benutzt, um die Lösungsstruktur des 39 kDa humanen Cdc37$_M$-Hsp90$_N$-Komplexes zu ermitteln. Weitere Details sind in Kapitel 7 zu finden.

Ergebnisse:

- Aufzeigen der Domänengrenzen, Reinigung und NMR-Bedingungen für Cdc37 (45 kDa)
- Komplette (99.5 %) Rückgratresonanzzuordnung der Hsp90-Bindungsdomäne von Cdc37
- Bestimmung der Struktur der humanen Hsp90-Bindungsdomäne von Cdc37 durch Röntgenkristallographie in Zusammenarbeit mit Prof. Dr. Roy Lancaster
- Abbildung der Wechselwirkungsgrenzfläche von Cdc37-Hsp90 durch Benutzung einer Vielzahl von NMR-Techniken, was als Beispiel für weitere Studien von großen Protein-Protein-Komplexen dient

- Erfolgreiche Nutzung von residualen dipolaren Kopplungen (RDC's) für eine genauere Strukturvorhersage des Protein-Protein-Komplexes

- Ermitteln der Struktur des humanen Cdc37-Hsp90 (40 kDa), durch Benutzung einer Kombination von NMR-Spektroskopie, Röntgenkristallographie und docking auf dem neuesten Stand der Technik

- PDB-Statistiken zeigen, dass es bis heute nur 49 Strukturen mit einem Molekular-gewicht über 39 kDa gibt, relativ zu 7728 durch NMR gelösten Strukturen. Die Struktur von Cdc37-Hsp90 ist eine der 49 Strukturen, die durch NMR gelöst wurde.

- Identifizierung des ‚Hot Spots' in der großen Wechselwirkungsgrenzfläche durch NMR und weitere Bestätigung durch eine Kombination aus Punktmutationen und NMR

Kapitel 8 beschreibt am Ende die NMR-Ergebnisse, dass Celastrol, ein kürzlich identifiziertes Triterpen, an Cdc37, nicht aber wie in der Literatur berichtet, an Hsp90 bindet und den Cdc37-Hsp90-Komplex zerstört. Weiterhin wird der molekulare Inhibierungsmechnismus von Cdc37, gelöst durch NMR, aufgezeigt.

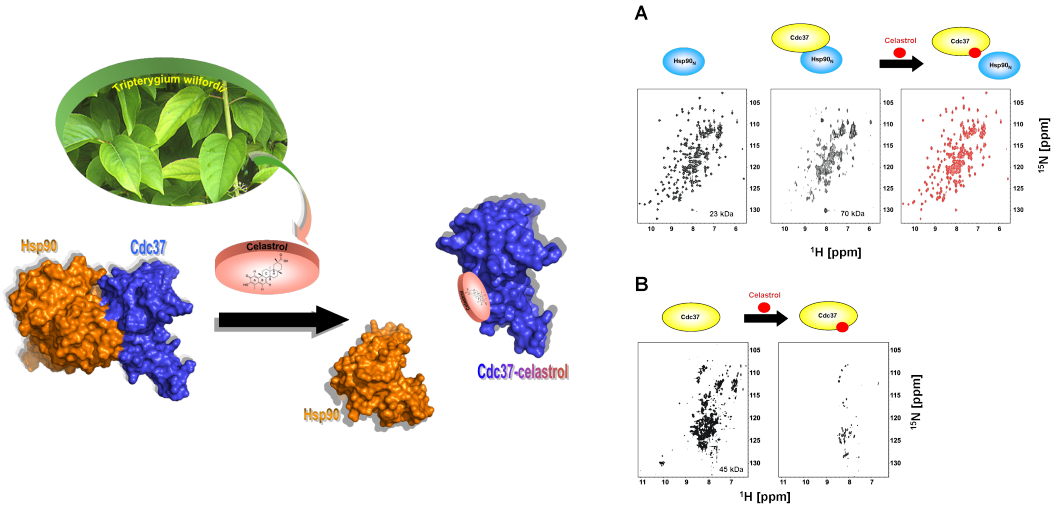


Abbildung 8.3: Celastrol bindet an Cdc37. Das „Cell division cycle"-Protein 37 (Cdc37) und das „Heat shock"-Protein (Hsp90) sind molekulare Chaperone, die für die Faltung und die Stabilisierung der Proteinkinasen, inklusive der onkogenen Kinasen, entscheidend sind. Hier zeigen wir durch NMR-Spektroskopie, dass Celastrol, ein kürzlich identifiziertes Triterpen nicht wie in der Literatur berichtet an Hsp90, sondern an Cdc37 bindet und somit den Cdc37-Hsp90-Komplex zerstört. Celastrol inaktiviert Cdc37 durch einen Thiol-vermittelten Mechanismus. Weitere Details sind in Kapitel 8 zu finden.

Ergebnisse:

- Nachweis, dass Cdc37 das richtige Ziel für Celastrol innerhalb des Protein-Protein-Komplexes (Cdc37-Hsp90) ist.
- Etablierung des exakten Mechanismus der Interaktion zwischen Celastrol und Cdc37 durch NMR
- Feststellung, dass die N-terminale und die mittlere Domäne von Cdc37 für die Bindung von Celastrol verantwortlich sind
- Diese Untersuchungen weisen die weitere Entwicklung von Inhibitoren zum richtigen molekularen Target innerhalb der Protein-Protein-Komplexe.

PUBLIKATIONSLISTE:

1. Folding and activity of cAMP-dependent protein kinase mutants: Thomas Langer, Sridhar Sreeramulu, Martin Vogtherr, Bettina Elshorst, Marco Betz, Ulrich Schieborr, Krishna Saxena, Harald Schwalbe, *FEBS Letters*, 2005, **579**, 4049-4054.
2. ^{1}H, ^{13}C and ^{15}N backbone resonance assignment of the Hsp90 binding domain of human Cdc37: Sridhar Sreeramulu, Jitendra Kumar, Christian Richter, Martin Vogtherr, Krishna Saxena, Thomas Langer, Harald Schwalbe, *Journal of Biomolecular NMR*, 2005, **32**, 262.
3. The Human Cdc37.Hsp90 Complex Studied by Heteronuclear NMR Spectroscopy: Sridhar Sreeramulu, Hendrik Jonker, Thomas Langer, Christian Richter, Roy Lancaster, Harald Schwalbe, *Journal of Biological Chemistry*, 2009, **284**, 3885-3896.
4. Molecular Mechanism of Inhibition of the Human Protein complex Hsp90-Cdc37, a Kinome Chaperone-Cochaperone, by Triterpene Celastrol, Sridhar Sreeramulu, Santosh Lakshmi Gande, Michael Göbel, Harald Schwalbe, 2009, *Angewandte Chemie International Edition English*, 2009, in press.

Bibliography

Abraham, D. J., Burger, A., and Lewis, F. L. (2003). Molecular biology of cancer. In Donald J. Abraham, editor, *Burger's medicinal chemistry and drug discovery. Vol. 5, Chemotherapeutic agents.* pages 1-50. Wiley Canada 6th edition. 19

Acton, T. B., Gunsalus, K. C., Xiao, R., Ma, L. C., Aramini, J., Baran, M. C., Chiang, Y.-W., Climent, T., Cooper, B., Denissova, N. G., Douglas, S. M., Everett, J. K., Ho, C. K., Macapagal, D., Rajan, P. K., Shastry, R., Shih, L.-Y., Swapna, G. V. T., Wilson, M., Wu, M., Gerstein, M., Inouye, M., Hunt, J. F., and Montelione, G. T. (2005). Robotic cloning and protein production platform of the northeast structural genomics consortium. Meth. Enzymol. *394*, 210-243. 12

Ahn, H.-C., Le, Y. T. H., Nagchowdhuri, P. S., Derose, E. F., Putnam-Evans, C., London, R. E., Markley, J. L., and Lim, K. H. (2006). NMR characterizations of an amyloidogenic conformational ensemble of the PI3K SH3 domain. Protein Sci. *15*(11), 2552-2557. 30

Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell *92*(3), 291-294. 47, 53

Amezcua, C. A., Harper, S. M., Rutter, J., and Gardner, K. H. (2002). Structure and interactions of pas kinase N-terminal pas domain: model for intramolecular kinase regulation. Structure *10*(10), 1349-1361. 59

Arkin, M. (2005). Protein-protein interactions and cancer: small molecules going in for the kill. Curr. Opin. Chem. Biol. *9*(3), 317-324. 48, 50

Arkin, M. R. and Wells, J. A. (2004). Small-molecule inhibitors of protein-protein

interactions: progressing towards the dream. Nat. Rev. Drug Discov. *3*(4), 301-317. 48, 50

Arumugam, S., Hemme, C. L., Yoshida, N., Suzuki, K., Nagase, H., Berjanskii, M., Wu, B., and Doren, S. R. V. (1998). TIMP-1 contact sites and perturbations of stromelysin 1 mapped by NMR and a paramagnetic surface probe. Biochemistry *37*(27), 9650-9657. 68

Aue, W., Bartholdi, E., and Ernst, R. (1976). Two-dimensional spectroscopy. application to nuclear magnetic resonance. J. Chem. Phys. *64*(5), 2229-2246. 56

Auguin, D., Barthe, P., Augé-Sénégas, M.-T., Hoh, F., Noguchi, M., and Roumestand, C. (2003). ^{1}H, ^{15}N and ^{13}C chemical shift assignments of the pleckstrin homology domain of the human protein kinase B (PKB/AKT). J. Biomol. NMR *27*(3), 287-288. 29

Auguin, D., Barthe, P., Augé-Sénégas, M.-T., Stern, M.-H., Noguchi, M., and Roumestand, C. (2004). Solution structure and backbone dynamics of the pleckstrin homology domain of the human protein kinase B (PKB/AKT). interaction with inositol phosphates. J. Biomol. NMR *28*(2), 137-155. 30

Basso, A. D., Solit, D. B., Chiosis, G., Giri, B., Tsichlis, P., and Rosen, N. (2002). AKT forms an intracellular complex with heat shock protein 90 (Hsp90) and Cdc37 and is destabilized by inhibitors of Hsp90 function. J. Biol. Chem. *277*(42), 39858-39866. 37

Bax, A., Clore, G., and Gronenborn, A. (1990a). ^{1}H^{1}H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^{1}H and ^{13}C spectra of ^{13}C-enriched proteins. J. Magn. Reson. *88*(2), 425-431. 56

Bax, A. and Ikura, M. (1991). An efficient 3D NMR technique for correlating the proton and ^{15}N backbone amide resonances with the α-carbon of the preceding residue in uniformly ^{15}N/^{13}C enriched proteins. J. Biomol. NMR *1*(1), 99-104. 56

Bax, A., Ikura, M., Kay, L., Torchia, D., and Tschudin, R. (1990b). Comparison of different modes of two-dimensional reverse-correlation NMR for the study of proteins. J. Magn. Reson. *86*(2), 304-318. 56

Bezsonova, I., Korzhnev, D. M., Prosser, R. S., Forman-Kay, J. D., and Kay, L. E. (2006). Hydration and packing along the folding pathway of SH3 domains by pressure-dependent NMR. Biochemistry *45*(15), 4711-4719. 30

Blume-Jensen, P. and Hunter, T. (2001). Oncogenic kinase signalling. Nature *411*(6835), 355-365. 28

Bocharov, E. V., Mayzel, M. L., Volynsky, P. E., Goncharuk, M. V., Ermolyuk, Y. S., Schulga, A. A., Artemenko, E. O., Efremov, R. G., and Arseniev, A. S. (2008). Spatial structure and pH-dependent conformational diversity of dimeric transmembrane domain of the receptor tyrosine kinase EphA1. J. Biol. Chem. *283*(43), 29385-29395. 30

Bodenhausen, G. and Ruben, D. (1980). Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. Chem. Phys. Lett. *69*(1), 185-189. 56

Bork, P., Jensen, L. J., von Mering, C., Ramani, A. K., Lee, I., and Marcotte, E. M. (2004). Protein interaction networks from yeast to human. Curr. Opin. Struct. Biol. *14*(3), 292-299. 48

Bowers, P. M., Strauss, C. E., and Baker, D. (2000). De novo protein structure determination using sparse NMR data. J. Biomol. NMR *18*(4), 311-318. 69

Braunschweiler, L. and Ernst, R. (1983). Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. J. Magn. Reson. *53*(3), 521-528. 56

Brünger, A. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr., Sect. D: Biol. Crystallogr. *54*(5), 905-921. 71

Brognard, J., Clark, A. S., Ni, Y., and Dennis, P. A. (2001). AKT/protein kinase B is constitutively active in non-small cell lung cancer cells and promotes cellular survival and resistance to chemotherapy and radiation. Cancer Res. *61*(10), 3986-3997. 1, 79

Brugge, J. S. (1986). Interaction of the rous sarcoma virus protein pp60src with the cellular proteins pp50 and pp90. Curr. Top. Microbiol. Immunol. *123*, 1-22. 37

Card, P. B. and Gardner, K. H. (2005). Identification and optimization of protein domains for NMR studies. Meth. Enzymol. *394*, 3-16. 58

Chen, G. and Goeddel, D. V. (2002). TNF-R1 signaling: a beautiful pathway. Science *296*(5573), 1634-1635. 20

Chène, P. (2006). Drugs targeting protein-protein interactions. ChemMedChem *1*(4), 400-411. 48

Choi, J., Chen, J., Schreiber, S. L., and Clardy, J. (1996). Structure of the FKBP12-rapamycin complex interacting with the binding domain of human frap. Science *273*(5272), 239-242. 50

Ciocca, D. R., Clark, G. M., Tandon, A. K., Fuqua, S. A., Welch, W. J., and McGuire, W. L. (1993). Heat shock protein Hsp70 in patients with axillary lymph node-negative breast cancer: prognostic implications. J. Natl. Cancer Inst. *85*(7), 570-574. 32

Clore, G. M. (2000). Accurate and rapid docking of protein-protein complexes on the basis of intermolecular nuclear overhauser enhancement data and dipolar couplings by rigid body minimization. Proc. Natl. Acad. Sci. U. S. A. *97*(16), 9021-9025. 68

Cochran, A. G. (2000). Antagonists of protein-protein interactions. Chem. Biol. *7*(4), R85-R94. 50

Cohen, F. E. and Prusiner, S. B. (1998). Pathologic conformations of prion proteins. Annu. Rev. Biochem. *67*, 793-819. 47

Cohen, S. L., Ferré-D'Amaré, A. R., Burley, S. K., and Chait, B. T. (1995). Probing the solution structure of the DNA-binding protein Max by a combination of proteolysis and mass spectrometry. Protein Sci. *4*(6), 1088-1099. 59

Conroy, S. E., Sasieni, P. D., Fentiman, I., and Latchman, D. S. (1998). Autoantibodies to the 90 kDa heat shock protein and poor survival in breast cancer patients. Eur. J. Cancer *34*(6), 942-943. 32

Croce, C. M. (2008). Oncogenes and cancer. N. Engl. J. Med. *358*(5), 502-511. 21

Deep, S., Im, S.-C., Zuiderweg, E. R. P., and Waskell, L. (2005). Characterization and calculation of a cytochrome c-cytochrome B5 complex using NMR data. Biochemistry *44*(31), 10654-10668. 65

Dominguez, C., Boelens, R., and Bonvin, A. M. J. J. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J. Am. Chem. Soc. *125*(7), 1731-1737. 70

Donaldson, L. W., Gish, G., Pawson, T., Kay, L. E., and Forman-Kay, J. D. (2002). Structure of a regulatory complex involving the Abl SH3 domain, the Crk SH2 domain, and a Crk-derived phosphopeptide. Proc. Natl. Acad. Sci. U. S. A. *99*(22), 14053-14058. 30

Downing, K. H. (2000). Structural basis for the interaction of tubulin with proteins and drugs that affect microtubule dynamics. Annu. Rev. Cell Dev. Biol. *16*, 89-111. 50

Dutta, R. and Inouye, M. (2000). GHKL, an emergent ATpase/kinase superfamily. Trends Biochem. Sci. *25*(1), 24-28. 35

Emerson, S. D., Madison, V. S., Palermo, R. E., Waugh, D. S., Scheffler, J. E., Tsao, K. L., Kiefer, S. E., Liu, S. P., and Fry, D. C. (1995). Solution structure of the Ras-binding domain of c-Raf-1 and identification of its Ras interaction surface. Biochemistry *34*(21), 6911-6918. 62

Esposito, V., Sjoberg, T., Das, R., Brown, S., Taylor, S. S., and Melacini, G. (2006). NMR assignment of the cAMP-binding domain a of the PKA regulatory subunit. J. Biomol. NMR *36 Suppl 1*, 64. 29

Fabbro, D., Ruetz, S., Buchdunger, E., Cowan-Jacob, S. W., Fendrich, G., Liebetanz, J., Mestan, J., O'Reilly, T., Traxler, P., Chaudhuri, B., Fretz, H., Zimmermann, J., Meyer, T., Caravatti, G., Furet, P., and Manley, P. W. (2002). Protein kinases as targets for anticancer agents: from inhibitors to useful drugs. Pharmacol. Ther. *93*(2-3), 79-98. 1, 79

Farmer II, B., Venters, R., Spicer, L., Wittekind, M., and Müller, L. (1992). A refocused and optimized HNCA: Increased sensitivity and resolution in large macromolecules. J. Biomol. NMR *2*(2), 195-202. 56

Ferrara, N. (2001). Role of vascular endothelial growth factor in regulation of physiological angiogenesis. Am. J. Physiol. Cell Physiol. *280*(6), C1358-C1366. 21

Ferrarini, M., Heltai, S., Zocchi, M. R., and Rugarli, C. (1992). Unusual expression and localization of heat-shock proteins in human tumor cells. Int. J. Cancer *51*(4), 613-619. 31

Fesik, S. W., Luly, J. R., Erickson, J. W., and Abad-Zapatero, C. (1988). Isotope-edited proton NMR study on the structure of a pepsin/inhibitor complex. Biochemistry *27*(22), 8297-8301. 61

Fiaux, J., Bertelsen, E. B., Horwich, A. L., and Wüthrich, K. (2002). NMR analysis of a 900k GroEL GroES complex. Nature *418*(6894), 207-211. 58

Fields, S. (2001). Proteomics. proteomics in genomeland. Science *291*(5507), 1221-1224. 7

Fischer, M. W., Losonczi, J. A., Weaver, J. L., and Prestegard, J. H. (1999). Domain orientation and dynamics in multidomain proteins from residual dipolar couplings. Biochemistry *38*(28), 9013-9022. 68

Folkman, J. (1971). Tumor angiogenesis: therapeutic implications. N. Engl. J. Med. *285*(21), 1182-1186. 21

Foster, M. P., Wuttke, D. S., Clemens, K. R., Jahnke, W., Radhakrishnan, I., Tennant, L., Reymond, M., Chung, J., and Wright, P. E. (1998). Chemical shift as a probe of molecular interfaces: NMR studies of DNA binding by the three amino-terminal zinc finger domains from transcription factor IIIA. J. Biomol. NMR *12*(1), 51-71. 62

Fry, D. C. and Vassilev, L. T. (2005). Targeting protein-protein interactions for cancer therapy. J. Mol. Med. *83*(12), 955-963. 48

Garrett, D. S., Seok, Y. J., Peterkofsky, A., Clore, G. M., and Gronenborn, A. M. (1997). Identification by NMR of the binding surface for the Histidine-containing phosphocarrier protein HPr on the N-terminal domain of enzyme I of the *Escherichia coli* phosphotransferase system. Biochemistry *36*(15), 4393-4398. 62

Garrett, D. S., Seok, Y. J., Peterkofsky, A., Gronenborn, A. M., and Clore, G. M. (1999). Solution structure of the 40,000 mr phosphoryl transfer complex between the N-terminal domain of enzyme I and Hpr. Nat. Struct. Biol. *6*(2), 166-173. 61

Gavin, A.-C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A.-M., Cruciat, C.-M., Remor, M., Höfert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.-A., Copley, R. R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., and Superti-Furga, G. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*(6868), 141-147. 47

Güntert, P. (2003). Automated NMR protein structure calculation. Prog. Nucl. Magn. Reson. Spectrosc. *43*(3-4), 105-125. 14

Gosser, Y. Q., Zheng, J., Overduin, M., Mayer, B. J., and Cowburn, D. (1995). The solution structure of Abl SH3, and its relationship to SH2 in the SH(32) construct. Structure *3*(10), 1075-1086. 29

Grammatikakis, N., Lin, J. H., Grammatikakis, A., Tsichlis, P. N., and Cochran, B. H. (1999). p50(Cdc37) acting in concert with Hsp90 is required for Raf-1 function. Mol. Cell. Biol. *19*(3), 1661-1672. 37, 41

Grammatikakis, N., Vultur, A., Ramana, C. V., Siganou, A., Schweinfest, C. W., Watson, D. K., and Raptis, L. (2002). The role of Hsp90N, a new member of the Hsp90 family, in signal transduction and neoplastic transformation. J. Biol. Chem. *277*(10), 8312-8320. 34

Gray, P. J., Prince, T., Cheng, J., Stevenson, M. A., and Calderwood, S. K. (2008). Targeting the oncogene and kinome chaperone Cdc37. Nat. Rev. Cancer *8*(7), 491-495. 35, 40, 42, 43, 44, 45, 51

Grzesiek, S. and Bax, A. (1992a). Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. J. Am. Chem. Soc. *114*(16), 6291-6293. 56

Grzesiek, S. and Bax, A. (1992b). An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. J. Magn. Reson. *99*(1), 201-207. 56

Grzesiek, S. and Bax, A. (1992c). Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. J. Magn. Reson. *96*(2), 432-440. 56

Grzesiek, S. and Bax, A. (1993). Amino acid type determination in the sequential assignment procedure of uniformly $^{13}C/^{15}N$-enriched proteins. J. Biomol. NMR *3*(2), 185-204. 56

Grzesiek, S., Stahl, S., Wingfield, P., and Bax, A. (1996). The CD4 determinant for downregulation by HIV-1 nef directly binds to Nef. mapping of the Nef binding surface by NMR. Biochemistry *35*(32), 10256-10261. 62

Hanada, M., Feng, J., and Hemmings, B. A. (2004). Structure, regulation and function of PKB/AKT-a major therapeutic target. Biochim. Biophys. Acta *1697*(1-2), 3-16. 1, 79

Hanks, S. K. and Hunter, T. (1995). Protein kinases 6. the Eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. FASEB J. *9*(8), 576-596. 22

Hansen, M. R., Mueller, L., and Pardi, A. (1998). Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. Nat. Struct. Biol. *5*(12), 1065-1074. 68

Harris, M. B., Bartoli, M., Sood, S. G., Matts, R. L., and Venema, R. C. (2006). Direct interaction of the cell division cycle 37 homolog inhibits endothelial nitric oxide synthase activity. Circ. Res. *98*(3), 335-341. 38

Hennings, H., Glick, A. B., Greenhalgh, D. A., Morgan, D. L., Strickland, J. E., Tennenbaum, T., and Yuspa, S. H. (1993). Critical aspects of initiation, promotion, and progression in multistage epidermal carcinogenesis. Proc. Soc. Exp. Biol. Med. *202*(1), 1-8. 19

Henzler-Wildman, K. and Kern, D. (2007). Dynamic personalities of proteins. Nature *450*(7172), 964-972. 55

Hieronymus, H., Lamb, J., Ross, K. N., Peng, X. P., Clement, C., Rodina, A., Nieto, M., Du, J., Stegmaier, K., Raj, S. M., Maloney, K. N., Clardy, J., Hahn, W. C., Chiosis, G., and Golub, T. R. (2006). Gene expression signature-based chemical genomic prediction identifies a novel class of Hsp90 pathway modulators. Cancer Cell *10*(4), 321-330. 45, 51

Hiroaki, H., Klaus, W., and Senn, H. (1996). Determination of the solution structure of the SH3 domain of human p56 Lck tyrosine kinase. J. Biomol. NMR *8*(2), 105-122. 29

Honndorf, V. S., Coudevylle, N., Laufer, S., Becker, S., and Griesinger, C. (2008). Dynamics in the p38α MAP kinase-SB203580 complex observed by liquid-state NMR spectroscopy. Angew. Chem., Int. Ed. *47*(19), 3548-3551. 30

Horst, R., Bertelsen, E. B., Fiaux, J., Wider, G., Horwich, A. L., and Wüthrich, K. (2005). Direct NMR observation of a substrate protein bound to the chaperonin GroEL. Proc. Natl. Acad. Sci. U. S. A. *102*(36), 12748-12753. 58

Huang, Y. J., Powers, R., and Montelione, G. T. (2005). Protein NMR recall, precision, and F-measure scores (RPF scores): structure quality assessment measures based on information retrieval statistics. J. Am. Chem. Soc. *127*(6), 1665-1674. 14

Huang, Y. J., Tejero, R., Powers, R., and Montelione, G. T. (2006). A topology-constrained distance network algorithm for protein structure determination from NOESY data. Proteins *62*(3), 587-603. 14

Hunter, T. and Poon, R. Y. (1997). Cdc37: a protein kinase chaperone? Trends Cell Biol. *7*(4), 157-161. 37

Huse, M. and Kuriyan, J. (2002). The conformational plasticity of protein kinases. Cell *109*(3), 275-282. 27

Jahnke, W., Rüdisser, S., and Zurini, M. (2001). Spin label enhanced NMR screening. J. Am. Chem. Soc. *123*(13), 3149-3150. 68

Jardetzky, O. and Jardetzky, C. (1957). An interpretation of the proton magnetic resonance spectrum of Ribonuclease. J. Am. Chem. Soc. *79*(19), 5322-5323. 56

Jarymowycz, V. A. and Stone, M. J. (2006). Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. Chem Rev *106*(5), 1624-1671. 29

Jeener, J., Meier, B., Bachmann, P., and Ernst, R. (1979). Investigation of exchange processes by two-dimensional NMR spectroscopy. J. Chem. Phys. *71*(11), 4546-4553. 56

Jolly, C. and Morimoto, R. I. (2000). Role of the heat shock response and molecular chaperones in oncogenesis and cell death. J. Natl. Cancer Inst. *92*(19), 1564-1572. 31

Kaelin, W. G. (2005). The concept of synthetic lethality in the context of anticancer therapy. Nat. Rev. Cancer *5*(9), 689-698. 37

Kamal, A., Thao, L., Sensintaffar, J., Zhang, L., Boehm, M. F., Fritz, L. C., and Burrows, F. J. (2003). A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors. Nature *425*(6956), 407-410. 35, 36

Kaur, J. and Ralhan, R. (1995). Differential expression of 70 kDa heat shock-protein in human oral tumorigenesis. Int. J. Cancer *63*(6), 774-779. 32

Kay, L., Ikura, M., Tschudin, R., and Bax, A. (1990). Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. J. Magn. Reson. *89*(3), 496-514. 56

Kay, L. E. (1998). Protein dynamics from NMR. Nat. Struct. Mol. Biol. *5 Suppl*, 513-517. 28

Kay, L. E., Muhandiram, D. R., Farrow, N. A., Aubin, Y., and Forman-Kay, J. D. (1996). Correlation between dynamics and high affinity binding in an SH2 domain interaction. Biochemistry *35*(2), 361-368. 67

Kay, L. E., Muhandiram, D. R., Wolf, G., Shoelson, S. E., and Forman-Kay, J. D. (1998). Correlation between binding and dynamics at SH2 domain interfaces. Nat. Struct. Biol. *5*(2), 156-163. 67

Kigawa, T., Yabuki, T., Matsuda, N., Matsuda, T., Nakajima, R., Tanaka, A., and Yokoyama, S. (2004). Preparation of escherichia coli cell extract for highly productive cell-free protein expression. J. Struct. Funct. Genomics *5*(1-2), 63-68. 12

Kilby, P. M., Eldik, L. J. V., and Roberts, G. C. (1997). Identification of the binding site on S100B protein for the actin capping protein CapZ. Protein Sci. *6*(12), 2494-2503. 65

Kim, S. and Szyperski, T. (2003). GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral information. J. Am. Chem. Soc. *125*(5), 1385-1393. 14

Kimura, E., Enns, R. E., Alcaraz, J. E., Arboleda, J., Slamon, D. J., and Howell, S. B. (1993). Correlation of the survival of ovarian cancer patients with mRNA expression of the 60 kDa heat-shock protein Hsp60. J. Clin. Oncol. *11*(5), 891-898. 32

Knighton, D. R., Xuong, N. H., Taylor, S. S., and Sowadski, J. M. (1991). Crystallization studies of cAMP-dependent protein kinase. cocrystals of the catalytic subunit with a 20 amino acid residue peptide inhibitor and MgATP diffract to 3.0 Å resolution. J. Mol. Biol. *220*(2), 217-220. 23

Kobe, B., Guss, M., and Huber, T. (2008). Structural proteomics: high-throughput methods. Methods Mol. Biol. *426*, v-vi. 8

Kreishman-Deitrick, M., Egile, C., Hoyt, D. W., Ford, J. J., Li, R., and Rosen, M. K. (2003). NMR analysis of methyl groups at 100-500 kDa: model systems and Arp2/3 complex. Biochemistry *42*(28), 8579-8586. 58

Krogh, A., Brown, M., Mian, I. S., Sjölander, K., and Haussler, D. (1994). Hidden Markov Models in computational biology. applications to protein modeling. J. Mol. Biol. *235*(5), 1501-1531. 59

Krone, P. H. and Sass, J. B. (1994). Hsp90α and Hsp90β genes are present in the Zebrafish and are differentially regulated in developing embryos. Biochem. Biophys. Res. Commun. *204*(2), 746-752. 34

Lamphere, L., Fiore, F., Xu, X., Brizuela, L., Keezer, S., Sardet, C., Draetta, G. F., and Gyuris, J. (1997). Interaction between Cdc37 and CDK4 in human cells. Oncogene *14*(16), 1999-2004. 37

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., Weinstock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D. R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H. M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R. W., Federspiel, N. A., Abola, A. P., Proctor, M. J., Myers, R. M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D. R., Olson, M. V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G. A., Athanasiou, M., Schultz, R., Roe, B. A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W. R., de la Bastide, M., Dedhia, N., Blöcker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J. A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D. G., Burge, C. B., Cerutti, L., Chen, H. C., Church, D., Clamp, M., Copley, R. R., Doerks, T., Eddy, S. R., Eichler, E. E., Furey, T. S., Galagan, J., Gilbert, J. G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L. S., Jones, T. A.,

Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W. J., Kitts, P., Koonin, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., McLysaght, A., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., Ponting, C. P., Schuler, G., Schultz, J., Slater, G., Smit, A. F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y. I., Wolfe, K. H., Yang, S. P., Yeh, R. F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A., Patrinos, A., Morgan, M. J., de Jong, P., Catanese, J. J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y. J., Szustakowki, J., and Consortium, I. H. G. S. (2001). Initial sequencing and analysis of the human genome. Nature *409*(6822), 860-921. 7

Lane, A. N., Kelly, G., Ramos, A., and Frenkiel, T. A. (2001). Determining binding sites in protein-nucleic acid complexes by cross-saturation. J. Biomol. NMR *21*(2), 127-139. 65

Langer, T., Vogtherr, M., Elshorst, B., Betz, M., Schieborr, U., Saxena, K., and Schwalbe, H. (2004). NMR backbone assignment of a protein kinase catalytic domain by a combination of several approaches: application to the catalytic subunit of cAMP-dependent protein kinase. ChemBioChem *5*(11), 1508-1516. 23, 29

Lee, P., Rao, J., Fliss, A., Yang, E., Garrett, S., and Caplan, A. J. (2002). The Cdc37 protein kinase-binding domain is sufficient for protein kinase activity and cell viability. J. Cell Biol. *159*(6), 1051-1059. 1, 41, 80

Leppä, S. and Sistonen, L. (1997). Heat shock response-pathophysiological implications. Ann. Med. *29*(1), 73-78. 31

Lewis, L. D. (2006). Cancer pharmacotherapy: 21st century 'magic bullets' and changing paradigms. Br. J. Clin. Pharmacol. *62*(1), 1-4. 18

Loregian, A. and Palù, G. (2005). Disruption of protein-protein interactions: towards new targets for chemotherapy. J. Cell. Physiol. *204*(3), 750-762. 48, 49, 50

Lundstrom, K. (2007). Structural genomics and drug discovery. J. Cell. Mol. Med. *11*(2), 224-238. 12

MacLean, M. and Picard, D. (2003). Cdc37 goes beyond Hsp90 and kinases. Cell Stress Chaperones *8*(2), 114-119. 37

Mandal, A. K., Lee, P., Chen, J. A., Nillegoda, N., Heller, A., DiStasio, S., Oen, H., Victor, J., Nair, D. M., Brodsky, J. L., and Caplan, A. J. (2007). Cdc37 has

distinct roles in protein kinase quality control that protect nascent chains from degradation and promote posttranslational maturation. J. Cell Biol. *176*(3), 319-328. 37

Mayer, M. P., Prodromou, C., and Frydman, J. (2009). The Hsp90 mosaic: a picture emerges. Nat. Struct. Mol. Biol. *16*(1), 2-6. 43

McCoy, M. A. and Wyss, D. F. (2002). Structures of protein-protein complexes are docked using only NMR restraints from residual dipolar coupling and chemical shift perturbations. J. Am. Chem. Soc. *124*(10), 2104-2105. 69

McLaughlin, S. H., Ventouras, L.-A., Lobbezoo, B., and Jackson, S. E. (2004). Independent ATpase activity of Hsp90 subunits creates a flexible assembly platform. J. Mol. Biol. *344*(3), 813-826. 35

Meyer, P., Prodromou, C., Hu, B., Vaughan, C., Roe, S. M., Panaretou, B., Piper, P. W., and Pearl, L. H. (2003). Structural and functional analysis of the middle segment of Hsp90: implications for ATP hydrolysis and client protein and cochaperone interactions. Mol. Cell *11*(3), 647-658. 34, 35

Meyer, P., Prodromou, C., Liao, C., Hu, B., Roe, S. M., Vaughan, C. K., Vlasic, I., Panaretou, B., Piper, P. W., and Pearl, L. H. (2004). Structural basis for recruitment of the ATpase activator Aha1 to the Hsp90 chaperone machinery. EMBO J. *23*(3), 511-519. 35

Millet, O., Loria, J., Kroenke, C., Pons, M., and Palmer III, A. (2000). The static magnetic field dependence of chemical exchange linebroadening defines the NMR chemical shift time scale. J. Am. Chem. Soc. *122*(12), 2867-2877. 64

Miyata, Y. and Nishida, E. (2004). CK2 controls multiple protein kinases by phosphorylating a kinase-targeting molecular chaperone, Cdc37. Mol. Cell. Biol. *24*(9), 4065-4074. 41

Müller, L. (1979). Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence. J. Am. Chem. Soc. *101*(16), 4481-4484. 56

Moore, S. K., Kozak, C., Robinson, E. A., Ullrich, S. J., and Appella, E. (1989). Murine 86 and 84 kDa heat shock proteins, cDNA sequences, chromosome assignments, and evolutionary origins. J. Biol. Chem. *264*(10), 5343-5351. 34

Mort-Bontemps-Soret, M., Facca, C., and Faye, G. (2002). Physical interaction of Cdc28 with Cdc37 in saccharomyces cerevisiae. Mol. Genet. Genomics *267*(4), 447-458. 37

Mueller, G. A., Choy, W. Y., Yang, D., Forman-Kay, J. D., Venters, R. A., and Kay, L. E. (2000). Global folds of proteins with low densities of noes using residual dipolar couplings: application to the 370-residue maltodextrin-binding protein. J. Mol. Biol. *300*(1), 197-212. 68

Mueller, G. A., Smith, A. M., Chapman, M. D., Rule, G. S., and Benjamin, D. C. (2001). Hydrogen exchange nuclear magnetic resonance spectroscopy mapping of antibody epitopes on the house dust mite allergen Der p 2. J. Biol. Chem. *276*(12), 9359-9365. 67

Muhandiram, D. and Kay, L. (1994). Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. J. Magn. Reson. B *103*(3), 203-216. 56

Mulhern, T. D., To, C., and Cheng, H.-C. (2002). ^{1}H, ^{13}C and ^{15}N chemical shift assignments of the SH2 domain of the Csk homologous kinase. J. Biomol. NMR *24*(4), 363-364. 29

Murphy, J. M., Korzhnev, D. M., Ceccarelli, D. F., Briant, D. J., Zarrine-Afsar, A., Sicheri, F., Kay, L. E., and Pawson, T. (2007). Conformational instability of the MARK3 UBA domain compromises ubiquitin recognition and promotes interaction with the adjacent kinase domain. Proc. Natl. Acad. Sci. U. S. A. *104*(36), 14336-14341. 30

Nakamura, T., Takahashi, H., Takeuchi, K., Kohno, T., Wakamatsu, K., and Shimada, I. (2005). Direct determination of a membrane-peptide interface using the nuclear magnetic resonance cross-saturation method. Biophys. J. *89*(6), 4051-4055. 66

Nakanishi, T., Miyazawa, M., Sakakura, M., Terasawa, H., Takahashi, H., and Shimada, I. (2002). Determination of the interface of a large protein complex by transferred cross-saturation measurements. J. Mol. Biol. *318*(2), 245-249. 66

Nishida, N., Sumikawa, H., Sakakura, M., Shimba, N., Takahashi, H., Terasawa, H., Suzuki, E.-I., and Shimada, I. (2003). Collagen-binding mode of vWF-A3 domain determined by a transferred cross-saturation experiment. Nat. Struct. Biol. *10*(1), 53-58. 66

Noble, M. E. M., Endicott, J. A., and Johnson, L. N. (2004). Protein kinase inhibitors: insights into drug design from structure. Science *303*(5665), 1800-1805. 1, 79

Norwood, T., Boyd, J., Heritage, J., Soffe, N., and Campbell, I. (1990). Comparison of techniques for ^{1}H-detected heteronuclear ^{1}H^{15}N spectroscopy. J. Magn. Reson. *87*(3), 488-501. 56

Nowell, P. C. (1976). The clonal evolution of tumor cell populations. Science *194*(4260), 23-28. 19

Okishio, N., Tanaka, T., Fukuda, R., and Nagai, M. (2001). Role of the conserved acidic residue Asp21 in the structure of phosphatidylinositol 3-kinase Src homology 3 domain: circular dichroism and nuclear magnetic resonance studies. Biochemistry *40*(1), 119-129. 29

Olejniczak, E., Xu, R., and Fesik, S. (1992). A 4D HCCH-TOCSY experiment for assigning the side chain ^{1}H and ^{13}C resonances of proteins. J. Biomol. NMR *2*(6), 655-659. 56

Otting, G. (1993). Experimental NMR techniques for studies of protein-ligand interactions. Curr. Opin. Struct. Biol. *3*(5), 760-768. 62

Palmer, A. G., Kroenke, C. D., and Loria, J. P. (2001). Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. Meth. Enzymol. *339*, 204-238. 64

Panaretou, B., Siligardi, G., Meyer, P., Maloney, A., Sullivan, J. K., Singh, S., Millson, S. H., Clarke, P. A., Naaby-Hansen, S., Stein, R., Cramer, R., Mollapour, M., Workman, P., Piper, P. W., Pearl, L. H., and Prodromou, C. (2002). Activation of the ATpase activity of Hsp90 by the stress-regulated cochaperone Aha1. Mol. Cell *10*(6), 1307-1318. 35

Pascale, R. M., Simile, M. M., Calvisi, D. F., Frau, M., Muroni, M. R., Seddaiu, M. A., Daino, L., Muntoni, M. D., Miglio, M. R. D., Thorgeirsson, S. S., and Feo, F. (2005). Role of Hsp90, Cdc37, and CRM1 as modulators of P16(INK4A) activity in rat liver carcinogenesis and human liver cancer. Hepatology *42*(6), 1310-1319. 40

Pearl, L. H. (2005). Hsp90 and Cdc37 - a chaperone cancer conspiracy. Curr. Opin. Genet. Dev. *15*(1), 55-61. 1, 80

Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc. Natl. Acad. Sci. U. S. A. *94*(23), 12366-12371. 58

Pervushin, K., Vögeli, B., and Eletsky, A. (2002). Longitudinal (1)H relaxation optimization in TROSY NMR spectroscopy. J. Am. Chem. Soc. *124*(43), 12898-12902. 14

Picard, D. (2002). Heat-shock protein 90, a chaperone for folding and regulation. Cell. Mol. Life Sci. *59*(10), 1640-1648. 31

Pintar, A., Hensmann, M., Jumel, K., Pitkeathly, M., Harding, S. E., and Campbell, I. D. (1996). Solution studies of the SH2 domain from the Fyn tyrosine kinase: secondary structure, backbone dynamics and protein association. Eur. Biophys. J. *24*(6), 371-380. 29

Prince, T., Sun, L., and Matts, R. L. (2005). Cdk2: a genuine protein kinase client of Hsp90 and Cdc37. Biochemistry *44*(46), 15287-15295. 40

Prodromou, C., Roe, S. M., O'Brien, R., Ladbury, J. E., Piper, P. W., and Pearl, L. H. (1997). Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. Cell *90*(1), 65-75. 34

Prodromou, C., Siligardi, G., O'Brien, R., Woolfson, D. N., Regan, L., Panaretou, B., Ladbury, J. E., Piper, P. W., and Pearl, L. H. (1999). Regulation of Hsp90 ATpase activity by tetratricopeptide repeat (TPR)-domain co-chaperones. EMBO J. *18*(3), 754-762. 35

Qin, J., Vinogradova, O., and Gronenborn, A. M. (2001). Protein-protein interactions probed by nuclear magnetic resonance spectroscopy. Meth. Enzymol. *339*, 377-389. 62

Ralhan, R. and Kaur, J. (1995). Differential expression of Mr 70,000 heat shock protein in normal, premalignant, and malignant human uterine cervix. Clin. Cancer Res. *1*(10), 1217-1222. 32

Ran, X., Qin, H., Liu, J., Fan, J.-S., Shi, J., and Song, J. (2008). NMR structure and dynamics of human ephrin-B2 ectodomain: the functionally critical C-D and G-H loops are highly dynamic in solution. Proteins *72*(3), 1019-1029. 29

Reed, S. I. (1980). The selection of *s. cerevisiae* mutants defective in the 'start' event of cell division. Genetics *95*(3), 561-577. 20, 37

Reed, S. I., de Barros Lopes, M. A., Ferguson, J., Hadwiger, J. A., Ho, J. Y., Horwitz, R., Jones, C. A., Lörincz, A. T., Mendenhall, M. D., and Peterson, T. A. (1985). Genetic and molecular analysis of division control in yeast. Cold Spring Harb. Symp. Quant. Biol. *50*, 627-634. 37

Riek, R., Wider, G., Pervushin, K., and Wüthrich, K. (1999). Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. Proc. Natl. Acad. Sci. U. S. A. *96*(9), 4918-4923. 58

Roe, S. M., Ali, M. M. U., Meyer, P., Vaughan, C. K., Panaretou, B., Piper, P. W., Prodromou, C., and Pearl, L. H. (2004). The mechanism of Hsp90 regulation by the protein kinase-specific cochaperone p50(Cdc37). Cell *116*(1), 87-98. 35, 41

Salomon, D. S., Brandt, R., Ciardiello, F., and Normanno, N. (1995). Epidermal growth factor-related peptides and their receptors in human malignancies. Crit. Rev. Oncol. Hematol. *19*(3), 183-232. 19

Santarosa, M., Favaro, D., Quaia, M., and Galligioni, E. (1997). Expression of heat shock protein 72 in renal cell carcinoma: possible role and prognostic implications in cancer patients. Eur. J. Cancer *33*(6), 873-877. 32

Scherf, T., Hiller, R., and Anglister, J. (1995). NMR observation of interactions in the combining site region of an antibody using a spin-labeled peptide antigen and NOESY difference spectroscopy. FASEB J. *9*(1), 120-126. 68

Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., Hartl, F. U., and Moarefi, I. (2000). Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. Cell *101*(2), 199-210. 35

Schnell, J. R., Zhou, G.-P., Zweckstetter, M., Rigby, A. C., and Chou, J. J. (2005). Rapid and accurate structure determination of coiled-coil domains using NMR dipolar couplings: application to cGMP-dependent protein kinase iα. Protein Sci. *14*(9), 2421-2428. 30

Scholz, G., Hartson, S. D., Cartledge, K., Hall, N., Shao, J., Dunn, A. R., and Matts, R. L. (2000). p50(Cdc37) can buffer the temperature-sensitive properties of a mutant of Hck. Mol. Cell. Biol. *20*(18), 6984-6995. 41

Schwarze, S. R., Fu, V. X., and Jarrard, D. F. (2003). Cdc37 enhances proliferation and is necessary for normal human prostate epithelial cell survival. Cancer Res. *63*(15), 4614-4619. 37, 40

Schweimer, K., Kiessling, A., Bauer, F., Hör, S., Hoffmann, S., Rösch, P., and Sticht, H. (2003). Sequence-specific ^{1}H, ^{13}C and ^{15}N resonance assignments of the SH3-SH2 domain pair from the human tyrosine kinase Lck. J. Biomol. NMR *27*(4), 405-406. 29

Scott, A., Pantoja-Uceda, D., Koshiba, S., Inoue, M., Kigawa, T., Terada, T., Shirouzu, M., Tanaka, A., Sugano, S., Yokoyama, S., and Güntert, P. (2004). NMR assignment of the SH2 domain from the human feline sarcoma oncogene FES. J. Biomol. NMR *30*(4), 463-464. 29

Selkoe, D. J. (1998). The cell biology of β-amyloid precursor protein and presenilin in alzheimer's disease. Trends Cell Biol. *8*(11), 447-453. 47

Shao, J., Grammatikakis, N., Scroggins, B. T., Uma, S., Huang, W., Chen, J. J., Hartson, S. D., and Matts, R. L. (2001). Hsp90 regulates p50(Cdc37) function during the biogenesis of the activeconformation of the heme-regulated EIF2 α kinase. J. Biol. Chem. *276*(1), 206-214. 41

Shao, J., Irwin, A., Hartson, S. D., and Matts, R. L. (2003). Functional dissection of Cdc37: characterization of domain structure and amino acid residues critical for protein kinase binding. Biochemistry *42*(43), 12577-12588. 41

Sharma, S. K., Ramsey, T. M., and Bair, K. W. (2002). Protein-protein interactions: lessons learned. Curr. Med. Chem. Anticancer Agents *2*(2), 311-330. 48

Shen, Y., Atreya, H. S., Liu, G., and Szyperski, T. (2005). G-matrix Fourier Transform NOESY-based protocol for high-quality protein structure determination. J. Am. Chem. Soc. *127*(25), 9085-9099. 14

Siligardi, G., Panaretou, B., Meyer, P., Singh, S., Woolfson, D. N., Piper, P. W., Pearl, L. H., and Prodromou, C. (2002). Regulation of Hsp90 ATpase activity by the co-chaperone Cdc37p/p50Cdc37. J. Biol. Chem. *277*(23), 20151-20159. 41

Slattery, M. L., Samowitz, W., Ballard, L., Schaffer, D., Leppert, M., and Potter, J. D. (2001). A molecular variant of the APC gene at codon 1822: its association with diet, lifestyle, and risk of colon cancer. Cancer Res. *61*(3), 1000-1004. 19

Smith, J. R., Clarke, P. A., de Billy, E., and Workman, P. (2009). Silencing the cochaperone Cdc37 destabilizes kinase clients and sensitizes cancer cells to Hsp90 inhibitors. Oncogene *28*(2), 157-169. 35, 43

Smith, J. R. and Workman, P. (2009). Targeting Cdc37: an alternative, kinase-directed strategy for disruption of oncogenic chaperoning. Cell Cycle *8*(3), 362-372. 35, 43, 44, 45, 51

Sowadski, J. M. and Epstein, L. F. (2000). Protein kinases. In *Encyclopedia of life sciences* pages 1-13. John Wiley and Sons. 22, 23, 24

Spronk, C. A., Bonvin, A. M., Radha, P. K., Melacini, G., Boelens, R., and Kaptein, R. (1999). The solution structure of Lac repressor headpiece 62 complexed to a symmetrical Lac operator. Structure *7*(12), 1483-1492. 62

Sreedhar, A. S., Kalmár, E., Csermely, P., and Shen, Y.-F. (2004). Hsp90 isoforms: functions, expression and clinical importance. FEBS Lett. *562*(1-3), 11-15. 34

Stam, K., Heisterkamp, N., Grosveld, G., de Klein, A., Verma, R. S., Coleman, M., Dosik, H., and Groffen, J. (1985). Evidence of a new chimeric bcr/c-Abl mRNA in patients with chronic myelocytic leukemia and the Philadelphia chromosome. N. Engl. J. Med. *313*(23), 1429-1433. 20

Stebbins, C. E., Russo, A. A., Schneider, C., Rosen, N., Hartl, F. U., and Pavletich, N. P. (1997). Crystal structure of an Hsp90-geldanamycin complex: targeting of a protein chaperone by an antitumor agent. Cell *89*(2), 239-250. 34

Stepanova, L., Finegold, M., DeMayo, F., Schmidt, E. V., and Harper, J. W. (2000a). The oncoprotein kinase chaperone Cdc37 functions as an oncogene in mice and collaborates with both c-Myc and Cyclin D1 in transformation of multiple tissues. Mol. Cell. Biol. *20*(12), 4462-4473. 37, 40

Stepanova, L., Yang, G., DeMayo, F., Wheeler, T. M., Finegold, M., Thompson, T. C., and Harper, J. W. (2000b). Induction of human Cdc37 in prostate cancer correlates with the ability of targeted Cdc37 expression to promote prostatic hyperplasia. Oncogene *19*(18), 2186-2193. 40, 43

Söti, C., Rácz, A., and Csermely, P. (2002). A nucleotide-dependent molecular switch controls ATP binding at the C-terminal domain of Hsp90. N-terminal nucleotide binding unmasks a C-terminal binding pocket. J. Biol. Chem. *277*(9), 7066-7075. 35

Strauss, A., Bitsch, F., Cutting, B., Fendrich, G., Graff, P., Liebetanz, J., Zurini, M., and Jahnke, W. (2003). Amino-acid-type selective isotope labeling of proteins expressed in baculovirus-infected insect cells useful for NMR studies. J. Biomol. NMR *26*(4), 367-372. 29

Stuart, D. I., Jones, E. Y., Wilson, K. S., and Daenke, S. (2006). SPINE: Structural Proteomics IN Europe - the best of both worlds. Acta Crystallogr., Sect. D: Biol. Crystallogr. *62*(10). 10

Sussman, J. and Silman, I. (2003). Structural proteomics and its impact on the life sciences. World Scientific Hackensack, N.J. 8, 12

Tabb, D. L., McDonald, W. H., and Yates, J. R. (2002). Dtaselect and contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J. Proteome Res. *1*(1), 21-26. 12

Takahashi, H., Nakanishi, T., Kami, K., Arata, Y., and Shimada, I. (2000). A novel NMR method for determining the interfaces of large protein-protein complexes. Nat. Struct. Biol. *7*(3), 220-223. 64

Takeuchi, K., Yokogawa, M., Matsuda, T., Sugai, M., Kawano, S., Kohno, T., Nakamura, H., Takahashi, H., and Shimada, I. (2003). Structural basis of the Kcsa K(+) channel and agitoxin2 pore-blocking toxin interaction by using the transferred cross-saturation method. Structure *11*(11), 1381-1392. 66

Taylor, J. D., Fawaz, R. R., Ababou, A., Williams, M. A., and Ladbury, J. E. (2005). NMR assignment of the apo and peptide-bound SH2 domain from the rous sarcoma viral protein src. J. Biomol. NMR *32*(4), 339. 29

Teicher, B. (2000). Molecular targets and cancer therapeutics: discovery, development and clinical validation. Drug Resist. Updat. *3*(2), 67-73. 1, 79

Terasawa, K., Minami, M., and Minami, Y. (2005). Constantly updated knowledge of Hsp90. J. Biochem. *137*(4), 443-447. 31

Tjandra, N. and Bax, A. (1997). Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science *278*(5340), 1111-1114. 68

Tolman, J. R., Flanagan, J. M., Kennedy, M. A., and Prestegard, J. H. (1997). NMR evidence for slow collective motions in Cyanometmyoglobin. Nat. Struct. Biol. *4*(4), 292-297. 68

Toogood, P. L. (2002). Inhibition of protein-protein association by small molecules: approaches and progress. J. Med. Chem. *45*(8), 1543-1558. 48

Tugarinov, V., Choy, W.-Y., Orekhov, V. Y., and Kay, L. E. (2005). Solution NMR-derived global fold of a monomeric 82 kDa enzyme. Proc. Natl. Acad. Sci. U. S. A. *102*(3), 622-627. 58

Tugarinov, V. and Kay, L. E. (2004). An isotope labeling strategy for methyl TROSY spectroscopy. J. Biomol. NMR *28*(2), 165-172. 58

Tugarinov, V., Muhandiram, R., Ayed, A., and Kay, L. E. (2002). Four-dimensional NMR spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase G. J. Am. Chem. Soc. *124*(34), 10025-10035. 58

Ulmer, T. S., Werner, J. M., and Campbell, I. D. (2002). SH3-SH2 domain orientation in Src kinases: NMR studies of Fyn. Structure *10*(7), 901-911. 30

Vajpai, N., Strauss, A., Fendrich, G., Cowan-Jacob, S. W., Manley, P. W., Grzesiek, S., and Jahnke, W. (2008). Solution conformations and dynamics of Abl kinase-inhibitor complexes determined by NMR substantiate the different binding modes of imatinib/nilotinib and dasatinib. J. Biol. Chem. *283*(26), 18292-18302. 29, 30

Valay, J. G., Simon, M., Dubois, M. F., Bensaude, O., Facca, C., and Faye, G. (1995). The KIN28 gene is required both for rna polymerase ii mediated transcription and phosphorylation of the Rpb1p CTD. J. Mol. Biol. *249*(3), 535-544. 37

Vaughan, C. K., Gohlke, U., Sobott, F., Good, V. M., Ali, M. M. U., Prodromou, C., Robinson, C. V., Saibil, H. R., and Pearl, L. H. (2006). Structure of an Hsp90-Cdc37-Cdk4 complex. Mol. Cell *23*(5), 697-707. 41

Vaughan, C. K., Mollapour, M., Smith, J. R., Truman, A., Hu, B., Good, V. M., Panaretou, B., Neckers, L., Clarke, P. A., Workman, P., Piper, P. W., Prodromou, C., and Pearl, L. H. (2008). Hsp90-dependent activation of protein kinases is regulated by chaperone-targeted dephosphorylation of Cdc37. Mol. Cell *31*(6), 886-895. 41

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Miklos, G. L. G., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A.,

Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Francesco, V. D., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M. L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y. H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N. N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J. F., Guigó, R., Campbell, M. J., Sjolander, K. V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y. H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A., and Zhu, X. (2001). The sequence of the human genome. Science *291*(5507), 1304-1351. 7, 17, 18

Veverka, V., Lennie, G., Crabbe, T., Bird, I., Taylor, R. J., and Carr, M. D. (2006). NMR assignment of the mTor domain responsible for rapamycin binding. J. Biomol. NMR *36 Suppl 1*, 3. 29

Vinarov, D. A., Newman, C. L. L., and Markley, J. L. (2006). Wheat germ cell-free platform for eukaryotic protein production. FEBS J. *273*(18), 4160-4169. 12

Vogtherr, M., Saxena, K., Grimme, S., Betz, M., Schieborr, U., Pescatore, B., Langer, T., and Schwalbe, H. (2005). NMR backbone assignment of the mitogen-activated protein (MAP) kinase p38. J. Biomol. NMR *32*(2), 175. 29

Wang, C., Pawley, N. H., and Nicholson, L. K. (2001). The role of backbone motions in ligand binding to the c-Src SH3 domain. J. Mol. Biol. *313*(4), 873-887. 30

Wegele, H., Müller, L., and Buchner, J. (2004). Hsp70 and Hsp90-a relay team for protein folding. Rev. Physiol. Biochem. Pharmacol. *151*, 1-44. 31

Welch, W. J. and Feramisco, J. R. (1982). Purification of the major mammalian heat shock proteins. J. Biol. Chem. *257*(24), 14949-14959. 31

Wells, J. A. and McClendon, C. L. (2007). Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature *450*(7172), 1001-1009. 2, 47, 50, 80

White, A. W., Westwell, A. D., and Brahemi, G. (2008). Protein-protein interactions as targets for small-molecule therapeutics in cancer. Expert Rev. Mol. Med. *10*, e8. 48

Whitelaw, M. L., Hutchison, K., and Perdew, G. H. (1991). A 50 kDa cytosolic protein complexed with the 90 kDa heat shock protein (Hsp90) is the same protein complexed with pp60v-Src Hsp90 in cells transformed by the rous sarcoma virus. J. Biol. Chem. *266*(25), 16436-16440. 37

Whitesell, L. and Lindquist, S. L. (2005). Hsp90 and the chaperoning of cancer. Nat. Rev. Cancer *5*(10), 761-772. 1, 31, 32, 33, 34, 36, 80

Wiesner, S., Hantschel, O., Mackereth, C. D., Superti-Furga, G., and Sattler, M. (2005). NMR assignment reveals an α-helical fold for the F-actin binding domain of human Bcr-Abl/c-Abl. J. Biomol. NMR *32*(4), 335. 29

Wittekind, M. and Mueller, L. (1993). HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the α- and β-carbon resonances in proteins. J. Magn. Reson. B *101*(2), 201-205. 56

Yamazaki, T., Lee, W., Arrowsmith, C., Muhandiram, D., and Kay, L. (1994). A suite of triple resonance NMR experiments for the backbone assignment of ^{15}N, ^{13}C, 2h labeled proteins with high sensitivity. J. Am. Chem. Soc. *116*(26), 11655-11666. 56

Yang, S., Qu, S., Perez-Tores, M., Sawai, A., Rosen, N., Solit, D. B., and Arteaga, C. L. (2006). Association with Hsp90 inhibits Cbl-mediated down-regulation of mutant epidermal growth factor receptors. Cancer Res. *66*(14), 6990-6997. 45, 51

Yee, A., Gutmanas, A., and Arrowsmith, C. H. (2006). Solution NMR in structural genomics. Curr. Opin. Struct. Biol. *16*(5), 611-617. 12

Yin, H. and Hamilton, A. D. (2005). Strategies for targeting protein-protein interactions with synthetic agents. Angew. Chem., Int. Ed. *44*(27), 4130-4163. 48

Yokogawa, M., Takeuchi, K., and Shimada, I. (2005). Bead-linked proteoliposomes: a reconstitution method for NMR analyses of membrane protein-ligand interactions. J. Am. Chem. Soc. *127*(34), 12021-12027. 67

Yufu, Y., Nishimura, J., and Nawata, H. (1992). High constitutive expression of heat shock protein 90 α in human acute leukemia cells. Leuk. Res. *16*(6-7), 597-605. 31

Yun, B.-G. and Matts, R. L. (2005). Differential effects of Hsp90 inhibition on protein kinases regulating signal transduction pathways required for myoblast differentiation. Exp. Cell Res. *307*(1), 212-223. 38

Zamzami, N., Brenner, C., Marzo, I., Susin, S. A., and Kroemer, G. (1998). Subcellular and submitochondrial mode of action of Bcl-2-like oncoproteins. Oncogene *16*(17), 2265-2282. 20

Zhang, T., Hamza, A., Cao, X., Wang, B., Yu, S., Zhan, C.-G., and Sun, D. (2008a). A novel Hsp90 inhibitor to disrupt Hsp90/Cdc37 complex against pancreatic cancer cells. Mol. Cancer Ther. *7*(1), 162-170. 45, 51

Zhang, W., Hirshberg, M., McLaughlin, S. H., Lazar, G. A., Grossmann, J. G., Nielsen, P. R., Sobott, F., Robinson, C. V., Jackson, S. E., and Laue, E. D. (2004). Biochemical and structural studies of the interaction of Cdc37 with Hsp90. J. Mol. Biol. *340*(4), 891-907. 41, 42

Zhang, W., Smithgall, T. E., and Gmeiner, W. H. (1997). Three-dimensional structure of the Hck SH2 domain in solution. J. Biomol. NMR *10*(3), 263-272. 29

Zhang, Y., Oh, H., Burton, R. A., Burgner, J. W., Geahlen, R. L., and Post, C. B. (2008b). Tyr130 phosphorylation triggers syk release from antigen receptor by long-distance conformational uncoupling. Proc. Natl. Acad. Sci. U. S. A. *105*(33), 11760-11765. 30

Zheng, J., Trafny, E. A., Knighton, D. R., Xuong, N. H., Taylor, S. S., Eyck, L. F. T., and Sowadski, J. M. (1993). 2.2 a refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. Acta Crystallogr., Sect. D: Biol. Crystallogr. *49*(Pt 3), 362-365. 27

Zheng, L. and Lee, W.-H. (2001). The retinoblastoma gene: A prototypic and multifunctional tumor suppressor. Exp. Cell Res. *264*(1), 2-18. 19

Zuiderweg, E. R. P. (2002). Mapping protein-protein interactions in solution by NMR spectroscopy. Biochemistry *41*(1), 1-7. 61

Zweckstetter, M. and Bax, A. (2000). Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. J. Am. Chem. Soc. *122*(15), 3791-3792. 71

Acknowledgements

“ It’s a great chance to express my gratitude to everyone who helped me in their best possible ways... "

First of all, I would like to thank my boss, supervisor and “meinem Doktorvater” **Prof. Dr. Harald Schwalbe**, for accepting me as his student and mentoring me throughout my research. My deep sense of gratitude to him, for not only providing me the opportunity, but also nurtured me through this long journey of doctoral study. Dear Harald, I still remember the email you sent when I was in India, accepting me as your student, it was a joyful moment and much more joy comes today as I acknowledge that, at this very nice moment. You have provided me the freedom of thinking, encouraged and supported me in all aspects of research. Vielen Dank!

I would like to express my heartfelt and deep sense of gratitude to our secretary **Anna Paulus** who has done so much right from the day of my arrival to Frankfurt till date. My boss once called her as “Star of our group”, who else can realize better than myself. Dear Anna, thank you very much and I wish you all the very best.

This is the best chance to express my deepest gratitude to **Dr. Thomas Langer**, whose constant encouragement towards my projects, made me climb up the hill. Perhaps I carried the most stupid questions (research oriented!) and he was always available with a smile and ready to answer my questions. Dear Thomas, I thank you for everything and wish you all the best.

One person to whom I am deeply indebted is **Dr. Christian Richter**. He has had probably the hardest time in training me at the spectrometers, but patiently delivered all the necessary knowledge for which I am extremely grateful to him. Dear Christian I thank you very much and wish you all the very best.

Today this work would have probably not been so successful, had I not knocked at the doors of **Dr. Henry Jonker.** He was instrumental, helpful and insightful in giving the final shape to my projects. The useful tips I got from him, working at the spectrometers are commendable at this moment. Dear Henry, I thank you very much for all the help you have given and wish you all the very best. Needless to mention I liked your “Hertog Jan”!

At this moment I would also like to thank the support given by **Prof. Dr. Roy Lancaster** and **Prof. Dr. Michael Göbel**, while persuading my collaborative projects.

I would also like to thank, **Prof. Dr. Jens Wöhnert,** for his useful suggestions during my doctoral study. Dear Jens, I thank you for all the help you have rendered and wish you all the very best.

A special word of thanks to **Dr. Muruga Poopathi Raja** for the useful discussions I had with him during my study and also for going through the task of spell checking my thesis. Dear Raja, I thank you very much for all the help and wish you and your family, all the very best.

I wish to express my thanks to **Dr. Julia Wirmer-Bartoschek,** with whom I had the opportunity to work at the spectrometer. Dear Julia, I thank you for sharing your ideas and experience and wish you all the very best.

I am grateful to **Dr. Chandramouli Chillakuri**, **Dr. Sachin Badrinath Surade**, **Dr. Ankita Srivastava** and **Dr. Minhajuddin Sirajuddin,** for all their help during my crystallization work. Their useful suggestions and tips led me towards success. I thank and wish you all, the very best.

One of the stories of my research would have been incomplete without the continuous and patient efforts of **Dr. Ute Bahr**. Dear Dr. Ute Bahr, I thank you for all the help you have done and wish you the very best.

The famous middle room, needless to say has provided me the best time and I really enjoy being a part of it. My deep sense of thankfulness to **Dr. Karin Abarca Heidemann**, **Jitendra Kumar**, **Tanja Machnik**, **Dr. Jonas Noeske** und **Stephan Rehm**. I wish you all, all the very best!

There is one person, who always came forward and helped me, right from finding an apartment upon my arrival in Frankfurt, until the translation of my thesis summary "Zusammenfassung". Yes! He is **Stephan Rehm**. Dear Stephan, I thank you very much for all the help you have given me and wish you all the very best for your future endeavors. In future, if I get an opportunity, I will join you in cheering "Federer"!

My deepest gratitude to **Jitendra Kumar**, who has been a constant encouragement towards my work, without whom, I would not have learnt the art of problem solving attitude. I thank you very much and wish you all the very best.

I would not have gained more knowledge about proteins, had I not got an opportunity to work with **Tanja Machnik**. It was a great pleasure to work with you Tanja! I thank and wish you all the very best.

I would like my sincere thanks to **Daniel Mathieu**, **Friederike Heinicke** and **Christian Gerum** for their extensive help rendered during my work in the lab. I thank and wish you all, the very best.

It's certainly a nice occasion to thank the practicum students **Anna Schnurr**, **Jochen Stehle** and **Priyanka** who briefly worked with me, but did a great job by verifying my data and pushing the projects ahead. I thank and wish you all, the very best.

I would like to thank **Dr. Bettina Elshorst, Dr. Marco Betz, Dr. Ulrich Schieborr, Dr. Krishna Saxena,** and **Dr. Martin Vogtherr** for their timely help during my doctoral study.

I would like to extend my thanks to **Dr. Christian Richter**, **Dr. Henry Jonker**, **Prof. Dr. Jens Wöhnert** and **Dr. Johannes Gottfried Zimmermann**, for all the help at the spectrometers.

Big thanks to our tech support team **Dr. Elke Duchardt, Jan-Peter Ferner, Dr. Christian Schlörb, Martin Hähnke, Daniel Mathieu and Fabian Hiller**.

I would like extend my heartfelt thanks to all the present and past members of the AK Schwalbe for providing an excellent working atmosphere: **Dr. Karin Abarca Heidemann, Dr. Nuria Aboitiz, Dr. Katrin Ackermann, Prof. Hashim Al Hashimi, Dr. Aphrodite Anastasiadis-Pool, Marie Anders-Maurer, Tomislav Argirevic, Neda Bakhtiari, Katja Barthelmes, Dr. Holger Berk, Anna Bischoff, Janina Buck, Florian Buhr, Dr. Aleksejs Cerepanov, Dr. Emily Collins, Dr. Elke Duchardt, Jan-Peter Ferner, Dr. Boris Fürtig, Christian Gerum, Dr. Artur Gieldon, Dr. Jürgen Graf, Dr. Steffen Grimm, Martin Hähnke, Friederike Heinicke, Fabian Hiller, Dr. Serge Ilin, Dr. Henry Jonker, Ajit Paul Kaur, Melanie Koschinat, Jitendra Kumar, Anna Lena Lieblein, Tanja Machnik, Vijayalaxmi Manoharan, Daniel Mathieu, Sarah Mensch, Hillary Moberly, Dr. Muruga Poopathi Raja, Dr. Hamid Nasiri, Gerd Nielsen, Dr. Jonas Noeske, Senada Nozinovic, Anke Reining, Dr. Christian Richter, Jörg Rinnenthal, Dr. Kai Schlepckow, Dr. Christian Schlörb, Dr. Nicole Schmut, Robert Silvers, Stephan Rehm, Max Stadler, Jochen Stehle, Elke Stirnal, Anna Wacker, Dominic Wagner, Dr. Karla Werner, Dr. Julia Wirmer-Bartoschek, Prof. Jens Wöhnert, Dr. Johannes Gottfried Zimmermann.**

I would like to express my gratefulness to our secretaries **Anna Paulus** and **Elena Hartmann**, for their continued efforts in making all the official procedures successful. Dear Anna and Elena, I thank you both for all the help and wish, all the very best.

A special thanks to **Prof. Dr. Volker Dötsch** for agreeing to be my “Zweitgutachter”.

I cannot leave this platform of thanking, without mentioning my deepest heartfelt thanks to **Gisela Besser**, “meiner deutschen Oma”, in whose home I stayed for the past five years. Also is worth mentioning about *late* **Hubert Besser**, together with Gisela, gave me some basic German lessons to start with, in Frankfurt. I thank you both for the extensive support you have given me through out my stay in Frankfurt. Gisela, I wish you good health and all the very best in the coming years.

Finally, I would like to thank my family for their unconditional support and good wishes. Especially, I would like to thank my parents and my wife, for whom I dedicate this thesis.

Curriculum Vitae

Personal Data

Name:	Sridhar Sreeramulu
Date of Birth:	02. January 1975
Place of Birth:	Ambur (India)
Nationality:	Indian
Marital Status:	Married to Dr. Santosh Lakshmi Gande

School

03/79 – 07/89 Primary education (St. Anthony's English School) in Bangalore
03/89 – 04/92 High School (Vijaya High School) in Bangalore

University Education

09/92 – 09/94 Diploma in Pharmacy (Government College of Pharmacy) in Bangalore
09/95 – 09/99 Bachelor of Pharmacy (Government College of Pharmacy) in Bangalore
09/99 – 09/02 Master of Pharmacy (Government College of Pharmacy) in Bangalore
04/02 – 08/03 Junior Research fellow (Indian Institute of Science) in Bangalore
09/03 – 05/09 PhD student with Prof. Dr. Harald Schwalbe, Goethe-University, Frankfurt, Germany.

Academic Teachers

Prof. Dr. Raghavan Varadarajan, Prof. Dr. S. Narasimha Murthy, Prof. M. Lakshmana, Prof. M. S. Harish, Prof. T. S. Dharmarajan, Prof. M. D. Karvekar, Prof. M. S. Niranjan, Prof. V. Ramakrishna, Prof. K. L. Mariswamy, Prof. M. S. Srinath, Prof. N. G. Nanjundaswamy, Prof. S. Shashidhara, Prof. Uma. D. Murthy, Mr. J. Chandrashekar, Mr. G. Kalaskar, Mrs. B. P. Manjula, Mr. S. Shiva Shankar, Mr. Prabha Kumar, Mrs. Lakshmi, Mrs. Girija Bai, Mr. S. C. Nagaraj, Mr. K. T. Basava, Prof. Dr. Harald Schwalbe, Prof. Dr. Jens Wöhnert, Prof. Dr. Hashim Al-Hashimi, Dr. Christian Richter.