

TECHNISCHE UNIVERSITÄT MÜNCHEN
Max-Planck-Institut für Biochemie

**Structural and functional characterization
of human protein kinases
exemplified on the human
cyclin-dependent kinase 8/ cyclin C complex**

Elisabeth V. Schneider

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

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. M. Groll

Prüfer der Dissertation: 1. apl. Prof. Dr. Dr. h.c. R. Huber (i. R.)

2. Univ.-Prof. Dr. Dr. A. Bacher (i. R.)

Die Dissertation wurde am 27. 06. 2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 10. 09. 2012 angenommen.

Preface

Mein besonderer Dank gilt meinem Doktorvater *Herrn Prof. Dr. Robert Huber* für seine wertvollen und weitblickenden Ratschläge sowie für seine wohlwollende Unterstützung und Ermutigung.

Zudem danke ich *Herrn Dr. Torsten Neufeind* für die Chance bei Proteros Biostructures zu arbeiten und zu lernen. Die vorhandene sehr gute Ausstattung sowie die äußerst beeindruckende Kompetenz der Wissenschaftler half die Grundlage dieser Arbeit zu legen.

Insbesondere danke ich *Dr. Klaus Maskos*, der mir gewissenhaft sein großes Fachwissen vermittelt hat. Er hatte einen maßgeblichen Beitrag am Erfolg meiner Arbeit und hat mich stets geduldig bei Fragen und Problemen aller Art unterstützt; nicht zuletzt konnte er beim Publikationsschreiben immer den „Weg zum roten Faden“ aufzeigen.

Ebenso danke ich *Dr. Jark Böttcher* für die überaus kompetente Einführung in die Kristallographie und die angenehme wissenschaftliche Zusammenarbeit während der gesamten Doktorarbeit. Gemeinsam mit *Dr. Michael Blaesse* hat er wichtige Ideen zur unkonventionellen aber erfolgreichen Lösung der weltweit ersten CDK8/CycC Struktur beigetragen.

Dr. Lars Neumann verdanke ich das Erlernen und die „Entdeckung der Langsamkeit“ des Kinetic Screenings, was letztendlich in eine komplette SKR-Studie resultiert ist. Durch seine außerordentliche fachliche Kompetenz, seine Begeisterungsfähigkeit und seinen Führungsstil war die Zusammenarbeit stets sehr angenehm.

Stellvertretend für das Einlernen und die Unterstützung rund um die Bench-Arbeit in meiner Anfangszeit in den verschiedensten Abteilungen möchte ich mich bei allen ehemaligen und aktiven Mitgliedern der „*Klaus PP3*“ Gruppe im Bereich der Proteinproduktion sowie bei *Allegra Ritscher*, *Birgit Flicke* und *Stefanie Gspurning* im Screening-Bereich und *Sonja Lernert* und *Theresa Wolfram* (my chief crystallographic instructor) im Kristallographie-Bereich bedanken.

Auch *Dr. Sabine Höppner* war mir durch ihre CDK Expertise eine große Unterstützung rund um dieses Projekt.

Ganz besonders *Birgit Flicke* gebührt mein herzlicher Dank für ihre permanente kollegiale Unterstützung und ihren freundschaftlichen Zuspruch in all den schönen, lustigen und manchmal auch verrückten Momenten während dieser Doktorarbeit.

Insgesamt danke ich allen ehemaligen und aktiven Mitarbeitern von Proteros Biostructures für ein angenehmes und produktives Arbeitsklima und ihre Hilfsbereitschaft.

Mein großer Dank gebührt auch *Monika Schneider* für Ihre Unterstützung und Hilfe im organisatorischen Bereich rund um diese Doktorarbeit.

Einen sehr großen Rückhalt haben mir meine Freunde gegeben, vor allem *Haike*, die mit mir durch dick und dünn geht. Einen herzlichen Dank auch an *Steffen*, der zudem meinen Abbildungen noch den letzten graphischen Feinschliff gegeben hat. Stellvertretend seien *Conny, Susi, Claudi, Mino, Johann, Georg* und *Cornelia* genannt, die mir bei gemeinsamen Unternehmungen die Zeit bereichert haben und bei denen ich immer wieder Energie auftanken konnte.

Mein herzlicher Dank gilt ebenso *meinem Orchester* für all die schönen Augenblicke, die erfüllenden Konzerte und Konzertreisen - und nicht zu vergessen *Agatharied* und die Feiern. *Dididämm*.

Vor allem aber danke ich *meinen Eltern* für ihre lebenslange Unterstützung und dass sie mir all das ermöglicht haben, meiner Schwester *Michaela* für ihren permanenten Zuspruch und meiner Schwester *Susi* für wichtige Ratschläge.

Relevant publications

At the moment of submitting the present dissertation, a part of it was already published in a scientific journal while other articles are still under preparation.

Schneider, E. V., Böttcher, J., Blaesse, M., Neumann, L., Huber, R. & Maskos, K., (2011). The structure of CDK8/CycC implicates specificity in the CDK/cyclin family and reveals interaction with a deep pocket binder. *J Mol Biol.* **412**, 251-266.

Schneider, E. V., Böttcher, J., Huber, R., Maskos, K. & Neumann, L. (2012). The discovery of slowness in terms of compound binding kinetics: a Structure-Kinetic-Relationship study on CDK8/CycC. *manuscript submitted*.

List of Abbreviations

Å	angstrom, 1×10^{-10} m
ATP	adenosintriphosphate
CAK	CDK activating kinase
CBF	cyclin box fold
CDK	cyclin-dependent kinase
CTD	C-terminal domain
Cyc	cyclin
DNA	deoxyribonucleic acid
DNA pol II	DNA polymerase II
DTT	1,4-dithiothreitol
E. coli	Escherichia coli
EG	ethylene glycol
EK	enterokinase protease cleavage-site
GSH	reduced glutathione
GST	glutathione <i>S</i> -transferase
GTF	general transcription factor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
H _N	N-terminal helix (cyclin-specific nomenclature)
H _C	C-terminal helix (cyclin-specific nomenclature)
IMAC	immobilized metal ion affinity chromatography purification
INK4	inhibitor of CDK4 and CDK6
IPTG	isopropyl-β-D-thiogalactopyranoside
MAT1	ménage á trois
MES	2-(N-morpholino)ethanesulfonic acid
MOPS	3-(N-morpholino)propanesulfonic acid
MR	molecular replacement
OD(x)	optical density at a wavelength of x nm
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
PKA	protein kinase A
PMSF	phenylmethylsulfonyl fluoride
PRE	PreScission protease cleavage-site
P-TEFb	positive transcription elongation factor b
RB	retinoblastoma tumor suppressor protein

rmsd	root mean square deviation
RNA pol II	ribonucleotide acide polymerase II
SAR	structure-activity-relationship
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SKR	structure-kinetic-relationship
SLS	Swiss Light Source
STG	<i>S</i> -tag
TCEP	tris(2-carboxyethyl)phosphine
TEV	Tobacco Etch Virus protease
Tf	trigger factor
THB	thrombin protease
TLS	translation-liberation-screw-rotation
TRIS	tris(hydroxymethyl)aminomethane
TRX	thioredoxin
v/v	volume/volume
vCyclin	cyclin Herpesvirus Saimiri
w/v	weight/volume

Summary

The cyclin-dependent kinase 8 (CDK8)/Cyclin C (CycC) complex is a part of the Mediator of transcription that regulates mRNA transcription through the RNA polymerase II and was identified as a potent oncogene in colon cancerogenesis. Within this thesis, the determination of the first crystal structure of human CDK8/CycC in complex with the clinical relevant anticancer-drug sorafenib revealed the first small-molecule induced DFG-out conformation in the CDK-family. The overall association of the CDK8/CycC complex adopts an intermediate conformation between the binding angles observed for the classical cell cycle CDK2/cyclin A complex and the transcriptional CDK9/cyclin T complex. The additional CycC recognition helix at the N-terminus of CDK8 was shown to be involved in discrimination against other cyclins *in-vitro* through CDK8 truncations.

Subsequently a structure-kinetic-relationship (SKR) study described the dependencies between binding kinetics and compound structure specific for CDK8/CycC, whereby the change of the compounds' binding behavior from fast to slow binding kinetics is monitored and results in an optimized residence time. The flip of the DMG-motif to the inactive DMG-out kinase conformation does not affect the velocity of compound binding. Hydrogen-bonding with the kinase hinge-region seems to trigger slow compound binding which is quantified in terms of a low observed association rate constant and a low dissociation rate constant, but obviously has less impact on the extension of residence time than hydrophobic complementarity within the kinase front pocket. In sum, specific inhibition of CDK8/CycC and discrimination against isolated CDK8 could be achieved.

Moreover structural evidence is presented that questions the general phosphorylation-dependent two step model of CDK activation via a phosphorylation-independent activation of CDK8 through association with CycC.

Zusammenfassung

Der humane Cyclin-abhängige Kinase 8 (CDK8)/ Cyclin C (CycC) Komplex ist Teil des sogenannten Mediator Komplexes, welcher die Transkription durch die RNA-Polymerase II reguliert. CDK8 wurde als starkes Onkogen innerhalb der Colon-Cancerogenese identifiziert. In dieser Doktorarbeit wurde die erste CDK8/CycC-Kristallstruktur aufgeklärt, welche im Komplex mit dem klinisch relevanten Krebsmedikament sorafenib die erste durch einen Inhibitor induzierte DFG-out Konformation in der CDK-Familie enthüllte. Der CDK8/CycC Komplex nimmt eine intermediäre Konformation zwischen den Bindewinkeln des klassischen Zellzyklus Komplexes CDK2/Cyclin A und des transkriptionell aktiven CDK9/CycT Komplexes an. Die entscheidende Rolle der CycC-Erkennungshelix am N-Terminus von CDK8 für die Assoziation mit CycC, welche die Unterscheidung von Cyclinen ermöglicht, konnte durch *in-vitro* Deletionsversuche untermauert werden.

Eine Studie mit CDK8/CycC-spezifischen Inhibitoren aus einer Fragmentbibliothek zeigte die Beziehung zwischen deren Struktur und Bindekinetik auf. Dabei konnten die strukturellen Determinanten für die Veränderung der Bindekinetik der Inhibitoren von einer schnellen Bindekinetik zu einer langsamen mit einer optimierten Residenzzeit aufgezeigt werden. Die Konformationsänderung des DMG-Motifs in eine inaktive DMG-out Konformation hat hierbei offenbar keinerlei Einfluss auf die Bindegeschwindigkeit der Inhibitoren. Wasserstoffbrücken mit der sogenannten „hinge“ Region der Kinase scheint hierbei eine langsamere Dissoziation der chemischen Verbindung auszulösen, haben aber weniger Einfluss auf die Dauer der Residenzzeit als hydrophobe Wechselwirkungen in der sogenannten „front pocket“ der Kinase. Insgesamt konnten spezifische Inhibitoren von CDK8/CycC entwickelt werden und darüber hinaus auch Unterschiede im Vergleich zur Inhibition isolierter CDK8 herausgearbeitet werden.

Die vorhandenen Strukturen stellen zudem das allgemeine phosphorylierungsabhängige Zwei-Schritt-Modell der CDK-Aktivierung in Frage und weisen auf eine phosphorylierungsunabhängige Aktivierung durch die Assoziation mit CycC hin.

Index

1. Introduction	1
1.1 Cancerogenesis	1
1.2 The regulation of the cell cycle and the cell cycle CDKs	1
1.3 The transcriptional process and the involvement of the CTD kinases	3
1.4 The structural organization of the CDK/cyclin family and its implication on the CDK kinase function	5
1.5 CDK8/CycC is a prominent member of the CDK family due to its role in physiology and disease	9
1.6 Drug-ability of CDK8/CycC: type I and type II inhibitors	10
1.7 Aims of this thesis	13
2. Material and Methods	14
2.1 Molecular Cloning	14
2.1.1 Polymerase chain reaction (PCR) based DNA amplification and sub-cloning	14
2.1.2 Bacterial transformation	15
2.2 Site-directed mutagenesis	16
2.3 Bacterial expression	17
2.4 Insect cell culture and heterologous protein expression	17
2.5 Analysis of testexpressions	18
2.6 Sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE) analysis and protein staining	18
2.7 Western Blot Analysis	19
2.8 Determination of Protein and Nucleic Acid Concentration	19
2.9 HIS-tag/GST-tag purification	19
2.10 Size exclusion chromatography (SEC)	20
2.11 Proteins concentration steps	20
2.12 Purification Protocols	20
2.13 Crystallization, X-ray data collection, processing and refinement	21
2.14 CDK8/CycC K_d measurement	23
2.15 The Proteros Reporter Displacement Assay	24
2.16 Chemicals	26
3. Establishment of protein production, crystallization and first structure determination of the human CDK8/CycC/sorafenib complex	27
3.1 Testing the different CycC-constructs in <i>E. coli</i> shows that it is not a suitable expression system for CycC but reveals the most appropriate tag-fusion construct	27
3.2 Protein expression of CDK8 and CycC is only possible separately in SF9 cells	32
3.3 Establishment of the CDK8/CycC complex formation <i>in-vitro</i> and its purification	36
3.4 Crystallization and refinement of crystals of the CDK8/CycC/sorafenib complex	38
3.5 Determination of the crystal structure of the CDK8/CycC/sorafenib complex	40
4. The first structure of CDK8/CycC within the CDK8-DMG-out conformation implicates specificity within the CDK/Cyclin Family and reveals interaction with a deep pocket binder	44
4.1 The Overall Structure of CDK8/CycC	44
4.2 Crystal Structure of human CycC	46
4.3 Crystal structure of human CDK8	49

4.4	The conserved core of the CDK/cyclin interface of CDK8/CycC: comparison with CDK9/CycT1	52
4.5	The adaptable H _N helix of CycC is contacting CDK8	54
4.6	CDK8 contains an additional N-terminal CycC recognition helix	55
4.7	Sorafenib binds to the hinge region in CDK8 inducing DMG-out conformation	58
5.	The discovery of slowness in terms of compound binding kinetics: a Structure-Kinetic-Relationship study on CDK8/CycC	61
5.1	The primary screen and the discovery of slowness	61
5.2	Combining structural with kinetic data reveals that the flip of the CDK8 DMG-motif does not influence the kinetic binding behavior of the compounds of the SKR series	67
5.3	The X-ray structures of CDK8/CycC with slow and fast binding compounds indicate that the main contributions to residence time are conserved hydrogen-bonding and hydrophobic contacts within the CDK8 front pocket	75
5.4	Discussion	81
6.	CDK8/CycC in the DMG-in kinase conformation versus the CDK8/CycC DMG-out kinase conformation: the two faces of the CDK8/CycC complex	85
6.1	Binding assay to select potential CDK8/CycC inhibitors that target the CDK8 DMG-in kinase conformation	85
6.2	Structural data on compound 12 in the active site CDK8/CycC site	88
6.3	Arrangements of CDK8/CycC related to the DMG-in conformation	90
6.4	The active site of CDK8/CycC in absence of an inhibitor	93
6.5	The crucial influence of CycC binding on the CDK8 conformation allows for selective inhibition of CDK8/CycC and discrimination against the CDK8 conformation in absence of the cyclin	96
6.6	Comparison of the CycC dependence of further SKR compounds	99
7.	References	103
	Curriculum Vitae	113

1. Introduction

1.1 Cancerogenesis

According to the World Health Organization cancer is the major cause of death worldwide, compromising around 13% of all deaths. Thereby the main types affect tissues such as lung, stomach, liver, breast and the colorectal region. The complex process of tumorigenesis involves progressive malign transformation of normal human cells into tumor cells (Hanahan and Weinberg, 2000). Proteins mutated and deregulated and thus supporting the capability of cells to sustain cancerogenic conditions are termed oncogenes. Exemplarily major oncogenes/proto-oncogenes are found within the family of the gene coupled receptor proteins, the transcription factors, the proteases or the protein kinases, which constitute around 2% of the human genome (Rubin et al., 2000). In total six hallmarks of deregulated cell physiology mark manifestation of tumor development: independency of growth-signals, resistance against anti-growth signals, evasion of apoptosis (programmed cell death), infinite replicative potential, sustained angiogenesis and tissue invasion including metastasis (Hannahan and Weinberg, 2000). In general cells with a tumorigenic competence are pushed towards an uncontrolled, rather proliferative than quiescent state of cell cycle. Due to its involvement in the regulation of the cell division cycle, apoptosis, transcription and differentiation, the family of the cyclin-dependent kinases (CDK) is considered an attractive target for cancer therapy. CDK1, CDK2, CDK3, CDK4 and CDK6 are the basic players in the cell cycle whereas CDK7 appears to be involved in its regulation. At least the three C-terminal domain (CTD) kinases CDK7, CDK8 and CDK9 (Pinhero et al., 2004) as well as CDK10 and CDK11 are reported to be involved in the regulation of transcription while CDK5 is an important player in the neuronal development (Malumbres et al., 2005).

1.2 The regulation of the cell cycle and the cell cycle CDKs

Quiescent cells are within a resting state (G_0) and stop proliferating or trespass into the active proliferative cell cycle depending on extracellular growth-signals or terminal differentiation. In the proliferative cell cycle the phases of semi-conservative DNA replication (S) and chromosome segregation/mitosis (M) are separated by the gap phases (G_1 and G_2). The whole process of cell cycle progression is tightly controlled. The trespassing of a restriction point is a precondition for progression

from G_1 to S-phase. Thereby signaling of the retinoblastoma tumor suppressor protein (RB) plays an important role. RB binds a subset of the transcriptional factor E2F necessary for the induction of genes involved in cell cycle progression (reviewed by Weinberg, 1995). In its phosphorylated form pRB cannot bind E2F anymore thus enabling S-phase entry (Sherr, 1996).

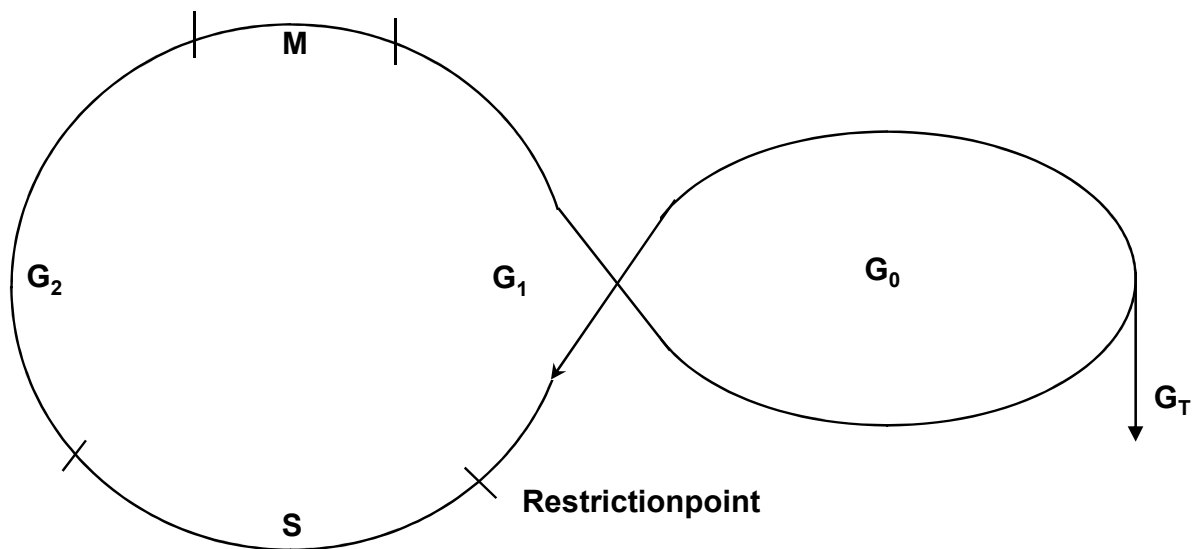


Figure 1 Cell Cycle

M = mitosis; S = semiconservative DNA replication; G_1 = gap 1 phase; G_2 = gap 2 phase; G_0 = resting state, G_T = terminal differentiation

In mammalian cells, the most important cell cycle CDKs are CDK1, CDK2, CDK3, CDK4 and CDK6. These are functionally dependent on specific association with their regulatory cyclin (Bregman et al., 2000; Dynlacht et al., 1997; Murray, 2004). CDK4/CycD, CDK6/CycD and CDK3/CycC (Sage, 2004; Ren et al., 2004) account for cell cycle progression over the restriction point from G_0 (quiescent state) to G_1 (proliferating cells) by phosphorylation of pRB (reviewed by Malumbres et al., 2005, Harper et al., 2001). In early G_1 -state, CDK2 in complex with CycE is believed to trigger complete phosphorylation of pRB and seems required for G_1 -S transition by licensing DNA origin of transcription. During progression of the S-phase the respective binding partner of CDK2 is CycA, which later associates with CDK1 as a co-regulator of S- G_2 and G_2 -M phases besides the CDK1/CycB complex. Other CDKs such as CDK11/CycL and CDK10 are supposed to be involved in mitosis but their exact involvement and role has to be investigated further (reviewed by Malumbres et al., 2005). In general the expression and degradation of the regulatory cyclin subunit of the cell cycle CDKs is tightly regulated by time-point specific events (reviewed by Viallard et al., 2001).

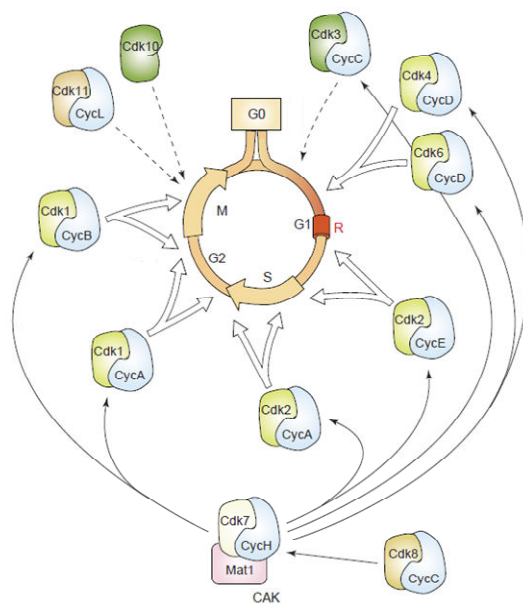


Figure 2 Cell cycle regulation by CDK/cyclin complexes

The discussed roles of the cell cycle CDK/cyclin pairs within the human cell cycle are schematically represented. CDK-activating kinase (CAK) has an additional role by activatory phosphorylation of the CDK/cyclin complexes, whose activity is in turn targeted by the CDK8/CycC pair. This figure was adapted from Malumbres et al., 2005.

1.3 The transcriptional process and the involvement of the CTD kinases

The transcriptional process or RNA synthesis to transcribe DNA nucleotide sequence information into RNA sequence information is catalyzed by the RNA polymerases. Thereby the RNA polymerase II (RNA pol II) is responsible to transcribe all protein-coding genes. Its structure, consisting of 12 subunits has been solved (Armache et al., 2005; Cramer et al., 2001) providing evidence that all RNA polymerases share a conserved and common transcription mechanism. During the process of RNA synthesis (initiation, elongation and termination) promoters are needed that consist of a core sequence element allowing binding of transcription factors, such as the TATA box. Besides RNA pol II controlling the process of transcription, basal general transcription factors (GTF: TFIIA, -B, -D, -E, -F, and -H) are needed for promoter recognition, promoter melting, abortive initiation and promoter escape (Kornberg, 2005). In general, transcription needs a tight regulation to avoid disadvantageous consequences for the organism such as malign transformation of tissues leading to cancerogenesis. Another essential player in the RNA pol II-mediated transcription machinery is the multi-protein assembly of the “Mediator of transcription” complex to stimulate basal transcription and regulate

eukaryotic mRNA synthesis (reviewed by Casamassimi et al., 2007) by both co-activation and co-repression. As the Mediator Complex is reported to be incapable of binding specific DNA sequences it exerts indirect influence by interacting with the C-terminal domain (CTD) of RNA pol II. The largest subunit of the CTD contains a 52-fold repeat of the heptapeptide $Y_1S_2P_3T_4S_5P_6S_7$ (Meinhart et al., 2005), with S_2 and S_5 being targets for regulatory reversible phosphoryl-transfer during the transcriptional cycle. Besides being a target for the Mediator of transcription the CTD is a platform for an assembly of factors that control transcription initiation, elongation, termination and mRNA processing (Hahn, 2004). In this respect, at least three CDK/cyclin pairs, the so-called “CTD-kinases” phosphorylate the CTD of the RNA pol II, thereby regulating transcription. The CTD-kinases, CDK7, CDK8 and CDK9 (Pinhero et al., 2004), associate with the so called C-type cyclins. This subfamily was identified as a more distantly related group of cyclins and seems to be constitutively expressed as the transcriptional cyclins do not show major fluctuation in abundance throughout the cell cycle in contrast to the classical cell cycle cyclins (Adamczewski et al., 1996, Garriga et al., 1998, Nigg et al., 1996, Rickert et al., 1996, Tassan et al., 1995). The CTD CDK/cyclin pairs all belong to multi-protein complexes. Of these, CDK8/CycC is known to belong to the Mediator of transcription as the “CDK submodule” together with MED12 and MED13 (reviewed by Casamassimi et al., 2007). CDK7/CycH associates with the protein “ménage à trois” (MAT1) to form the CDK activating kinase (CAK) that activates other CDKs by phosphorylation (Devault et al., 1995; Fisher et al., 1995; Harper and Elledge, 1998; Larochelle et al., 1998; Wallenfang and Seydoux, 2002). This trimeric complex is also part of the GTF TFII that is a main factor in transcription (Oelschlager, 2000). CDK9 associates with one of the cyclin T isoforms (T1, T2a or T2b) or cyclin K and other factors such as Bromodomain-containing protein 4 (Yang et al., 2005) to form the positive transcription elongation factor b (p-TEFb) and is involved in trespassing transcription from the initiation to the elongation phase (Baumli et al., 2010; Baumli et al., 2008; Fujinaga et al., 2004; Yamada et al., 2006; Marshall et al., 1996).

1.4 The structural organization of the CDK/cyclin family and its implication on the CDK kinase function

Several crystal structures of the family of the CDK/Cyclin have been solved. The best characterized CDK is CDK2, whose structure has been solved as monomer and in various complexes including the complex with the adaptor protein CKS1, the phosphatase KAP, the CDK inhibitor p27 or cyclin-bound recruitment peptides. Further available crystal structures are the various complexes such as human CDK4/CycD, CDK9/CycT1, CDK5/p25, CDK6/cyclin Herpesvirus Saimiri (vCyclin), CDK6/tumor suppressor inhibitor of CDK4 and CDK6 (INK4), CDK7, cyclin K, CycT/TAT/TAR, CycH and the yeast CycC analogue SRB11 (see Figure 3 and the detailed review by Lolli, 2010). The members of the CDK family all show the typical kinase-fold comprising N-lobe and C-lobe that are connected by the so-called hinge-region. The smaller N-lobe predominantly consists of β -sheets whereas the larger C-lobe predominantly forms a helical structure. ATP is bound in the catalytic cleft between the two lobes beneath the highly conserved glycine-rich (GXGX Φ G) β 1- β 2 loop, where Φ usually is tyrosine or phenylalanine as a “roof” of the phosphate-transfer. Moreover the activation loop (T-loop) serves as sterical barrier to block the active site from binding of both nucleotide and protein substrates in its closed and inactive conformation. Within the opened active kinase conformation a morphological change of the T-loop enables access for the substrates. The only conserved helix in the β -sheet rich N-lobe is the α C-helix. It contains an absolutely conserved glutamate (E91 in protein-kinase A (PKA)) that forms an ion pair with a lysine (K72 in PKA) to coordinate the α - and β phosphates of ATP (Bossemeyer et al., 1993; detailed review by Huse et al., 2002).

Ligands	Description	PDB code	Ligands	Description	PDB code
CDK2	Inactive, "closed" activation segment blocks access of substrate to active site, α C helix "out" misaligns residues involved in productive binding of ATP (see below)	1HCL	CDK6/KCyc/p18	Viral cyclin from Kaposi's sarcoma-associated herpes virus has a reduced interaction area when compared to CDK2/CycA. Reduced by distortion of CDK6 cyclin-binding surface by p18. CDK6 is in the inactive conformation observed in 1B17 and 1B18.	1G3N
CDK2	Asp145 pairs with Lys33 resulting in non-productive ATP binding	1HCK 1B38	CDK6/KCyc	Viral cyclin from herpes virus saimiri activates CDK6 as observed for CDK2 pushing "in" α C helix and moving activation segment in the "open" conformation. Interface area is particularly extended causing resistance to CN inhibition	1JOW
pThr160-CDK2	0.3% of full activity observed for pThr160-CDK2/CycA. Activation segment is flexible allowing access of substrate to the active site	1B39	CDK4/CycD3	Inactive CDK with α C helix "out" and "closed" activation segment. Reduced interface area compared to CDK2/CycA or CDK6/KCyc	3G33
CycA	CDK2 partner	1VIN	pThr172-CDK4/CycD1	Similar to 3G33 with higher flexibility observed for the phosphorylated activation segment	2W96
CDK2/CycA	1% of full activity observed for pThr160-CDK2/CycA. CycA does not show any structural change following CDK2 binding. In CDK2, α C helix is pushed "in". Glu51 from α C helix interacts with Lys33 replacing Asp145 that can now participate in correct orientation of ATP. Activation segment is "open" but substrate binding site is not preformed	1FIN	CDK5/p25	CDK in active conformation bound to the single cyclin box partner p25	1H4L
pThr160-CDK2/CycA	Fully active. Asp145 binds the Mn ion. pThr160 interacts with Arg50, Arg126 and Arg150 stabilizing the "open" activation segment.	1J5T	CycC	CDK8 partner	1ZP2
pThr160-CDK2/CycA	Consensus phosphorylation sequence S/T-P-X-K/R. Val164 adopts a left-handed conformation conferring proline-directed specificity. The trans proline directs the basic residue of the consensus sequence to the pThr160.	1OMZ	CycD	CDK9 partner	2I53
pThr160-CDK2/CycA	p27 binds to CycA recruitment site and to CDK2; inhibition is exerted by distorting CDK2 N-terminal lobe and by filling its catalytic cleft	1J5U	CycE2	CDK9 partner	2WVX
pThr160-CDK2/CycA	Peptides from E/F1, p53, p18, p27 and p107 bound to CycA recruitment site	1H24 1H25 1H26 1H27 1H28	pThr186-CDK9/CycT1	CDK in active conformation with α C helix "in" and "open" activation segment	3BLH
pThr160-CDK2/CycA	Substrate and recruitment peptide Mg/AMPPNP	2CCI	Mg/AMPPNP	Mg/AMPPNP in productive orientation as observed for CDK2/CycA in 1QMZ	3BLQ
pThr160-CDK2/CycA	30 residues CDK6 peptide with 7 residues in CDK2 active site, 12 residues bound to CycA and flexible 11 linking residues	2CIM	CycT1/TAT/TAR	Equine CycT1 in complex with TAT/TAR from equine infectious anemia virus (closely related to HIV). Interaction is centered on CycT1 H4/H5 loop	2WZH
pThr160-CDK2/CycA	pTyr15 causes steric blockade of substrate binding and favors non-productive ATP conformation	1FQ1	CycH	CDK7 partner	1KXU
pThr160-CDK2/KAP	KAP recognizes CDK2 through binding at its C-terminal lobe. CDK2 activation segment protrudes from the kinase body placing pThr160 inside KAP catalytic site	1BUH	pThr170-CDK7	Inactive, "closed" activation segment, α C helix "out"	1UA2
CDK2/CKS1	CKS1 binds at CDK2 C-terminal lobe and, forming an extended recognition surface close to CDK2 catalytic site, targets CDKs to other phosphoproteins	1W99	CDK2	Purvalanol B	1CKP
pThr160-CDK2/Cyc-E	CDK in active conformation. Greater interaction surface in respect to CycA through CycE specific insert in the H1/H2' loop	2B9R	CDK2	R-roscovitine	2A4L
CycB	CDK1 partner	2JGZ	CDK5/p25	R-roscovitine	1UUL
pThr160-CDK2/CycB	CDK in active conformation. Lower contact with CDK2 in respect to CycE and CycA with lower affinity as measured by SPR	1B17	CDK2	Deschloro-flavopiridol	Not deposited
CDK6/p16	CDK in inactive conformation with α C helix "out". Activation segment is differently located in respect to CDK2 and CDK2/CycA and is in the way of ATP binding. Inhibition by p16 exerted by distortion of the cyclin binding surface and distortion of ATP-binding site	1B18 1BLX	pThr186-CDK9/CycT1	Flavopiridol	3BLR
CDK6/p19	Similar to CDK6/p16 for structure 1B18. In 1BLX the activation segment is in the "closed" conformation similar to that observed in CDK2		CDK2	AT-7519 analogue	2VTT
			CDK2	BMS-387032	Not deposited
			CDK2	R-547	2PYD
			CDK6-KCyc	PD-0332991	2EUF

Figure 3 Overview of the most important crystal structures of the CDK/cyclin family

Figure adapted from Lolli, 2010

Similarly the family of the cyclins comprises a highly conserved cyclin box fold (CBF) that exhibits the common structural core of two canonical cyclin box folds, each consisting of five helices with additional helices at the N-terminus (H_N) as well as at the C-terminus (H_C) (Noble et al., 1997). However, in contrast to their CDK binding partner, non-conserved differences at a higher degree are found within the structures of the cyclins. As recently reviewed (Lolli, 2010 and Echaliier et al., 2010), the functional difference of the cyclins involved in either the control of the cell cycle (e.g. CycA, CycB, CycD, CycE) or the control of transcription (e. g. CycC, CycH, CycK, CycT) is reconciled within their structure. Thus the length of the above described CBF varies and differences of the orientation and number of the H_N and H_C helices are observable depending on the function of the cyclin. Thus the N-terminal H_N helix of the cell cycle cyclins is involved in recognition of their distinctive kinase with the H_N helices of CycA, CycB and CycE interacting with the CDK activation segment and the C-terminal lobe of CDK2. The H_N helix of the structurally less characterized transcriptional cyclins is located opposite the CDK binding surface and consequently cannot contribute to CDK recognition. Therefore the binding mode of a transcriptional CDK/cyclin pair differs from that of a transcriptional CDK/cyclin as exemplified in the crystal structure of the CDK9/CycT complex (Baumli et al., 2008) with a dramatically reduced binding surface as compared to CDK2/CycA (Lolli, 2010).

The common binding surface among all CDK/cyclin complexes is located in the CDK pre- α C-region (Figure 4, above) together with the α C helix and the post- α C region (Lolli, 2010; Echaliier et al., 2010). An analysis of the CDK/Cyclin interfaces of CDK2/CycA, CDK2/CycE, CDK9/CycT1 and CDK4/CycD revealed an extensively conserved feature for all hetero-dimers that consists of a hydrophobic network that shields an aromatic cyclin residue within its center from solvent (Echaliier et al., 2010). CDK/cyclin specificity *in-vivo* is achieved by only subtle changes within shape and electrostatic charge in terms of length, sequence and spatial orientation of the counteracting CDK/cyclin residues towards perfect complementarity of the CDK/cyclin pair formation of the CDK/cyclin pair (Lolli, 2010). However, the CDK/cyclin complex formation *in-vitro* is reported to be highly promiscuous based on the large common CDK binding surface on the cyclins (Lolli, 2010). Indeed it was postulated, that the same hydrophobic contacts are involved in mediating binding

surface between any cyclin with any CDK (Heitz et al., 1997). Exemplarily in the case of CycC, the cyclin is able to associate with both CDK8 and CDK3. Depending on the respective interacting CDK it then either exerts a function within the cell cycle (CDK3) or functions as transcriptional cyclin (CDK8).

A general mechanism of CDK activation and substrate specificity was based on the comprehensive structural information on CDK2 structures (Jeffrey et al., 1995; Pavletich, 1999, see Figure 4). First, the cyclin binds the corresponding CDK with the common binding surface located on the CDK α C helix. Thereby the α C-helix is rearranged towards the activation segment. The next step includes reversible phosphorylation of the CDK T-loop. The phospho-residue (T160p^{CDK2}) then interacts with a conserved arginine triad located within the α C helix/N-lobe, the T-loop and the C-lobe (R50^{CDK2}, R162^{CDK2} and R150^{CDK2}) within an electrostatic network. Thus the activation segment is maintained in an open conformation. Even though this activation mechanism has been regarded as generally valid for the CDK/cyclin family further structural studies of the CDK-family revealed evidence questioning this two-step activation model (Tarricone et al., 2001). A mechanism independent of activatory phosphorylation has been suggested for CDK5 complexed with p25, a part of CDK5-activating protein (Tarricone et al., 2001). Surprisingly in the structure of human CDK4/CycD association of the cyclin did not lead to an active kinase conformation (α C helix pushed in) (Day et al., 2009; Takaki et al., 2009).

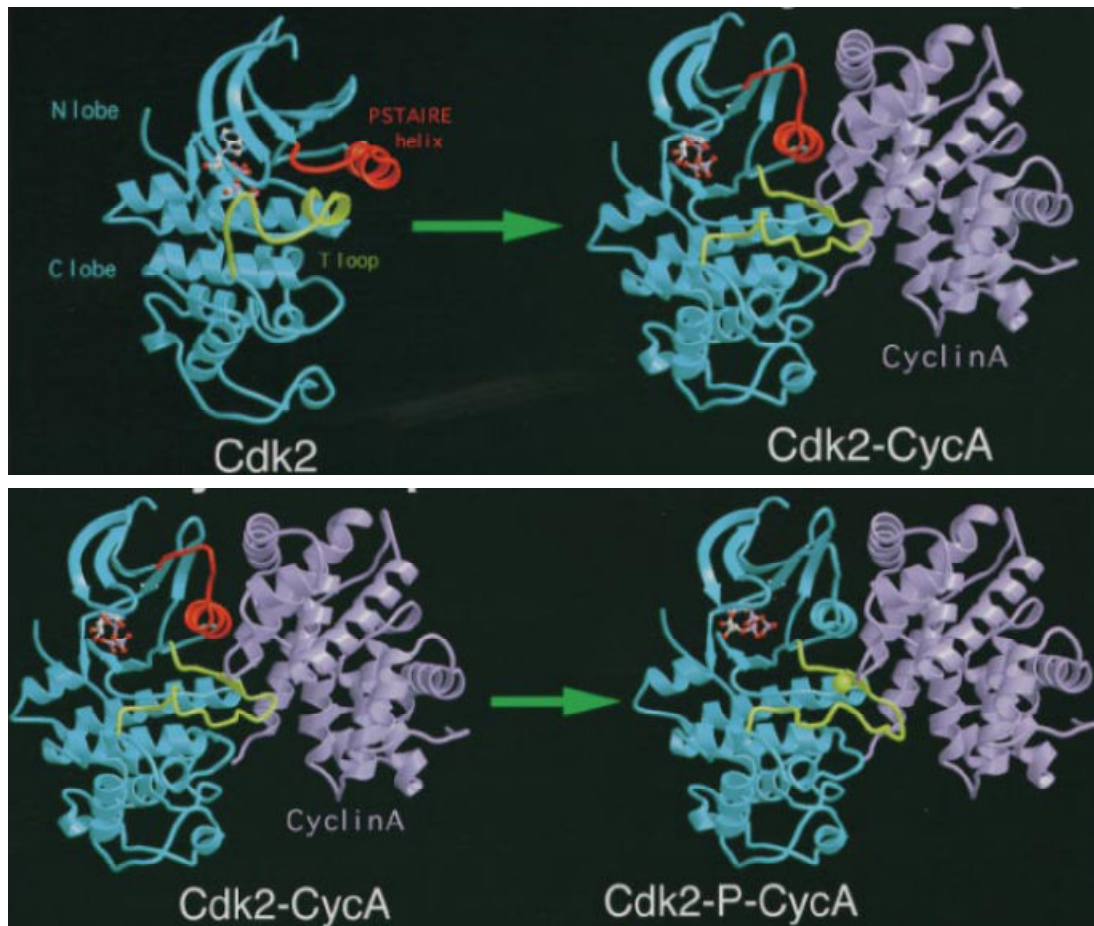


Figure 4 Two-step activation model of the CDK/cyclin family exemplified on CDK2/CycA

Figure adapted from Pavletich, 1999; *above* located illustration depicts the re-arrangement of the CDK2 α C helix towards the CDK2 activation segment by binding of CycA (step 1) and “melting” of a short helical turn within the CDK2 activation segment to enfold itself and relieve a sterical blockade. In this way E51^{CDK2} (α C helix) is positioned to interact with K33^{CDK2} to form the overall conserved salt-bridge that is essential within the kinase family. The second step of CDK kinase activation (*illustration below*) requires reversible phosphorylation of a CDK T-loop located S/T-residue. In CDK2 T160^{CDK2} becomes phosphorylated and interacts with three arginines located within the C-lobe, in the T-loop as well as in the N-lobe (R126^{CDK2}, R150^{CDK2} and R50^{CDK2}) thus leading to complete activation of CDK2, including complete catalytic competence, and maintaining the CDK2 activation within an open conformation.

1.5 CDK8/CycC is a prominent member of the CDK family due to its role in physiology and disease

Since the discovery of the CDK8/CycC pair (Liao et al., 1995; Tassan et al., 1995), its importance as a possible drug target has been discovered more and more. The CDK8/CycC analogue SRB10/SRB11 suppresses transcription both directly by phosphorylation of RNA pol II in yeast (Hengartner et al.,

1998). An indirect suppression was described for human CDK8/CycC by phosphorylation of human CycH (Akoulitchev et al., 2000). Human CDK8/CycC also is a positive regulator of transcriptional gene activity targeting for example p53 genes (Donner et al., 2007) and the serum response network (Donner et al., 2010) implying oncogenic effects. Besides, the CDK module was postulated to be involved in initiation and re-initiation of transcription (Knuesel et al., 2009a). A physiological role independent of the Mediator complex was described for CDK8/CycC/MED12 in its function as a histone kinase (Knuesel et al., 2009b). In the past years evidence evolved for an important role of the CDK8 kinase in colon cancerogenesis involving β -catenin/wnt signaling. Elevated levels of CDK8 were identified in colon cells to sustain cancerogenic conditions by means of β -catenin signaling (Morris et al., 2008) and the activity of CDK8 was necessary for β -catenin triggered malign transformation in colon cells when comparing over-expression of an active to a kinase-dead CDK8 mutant (Firestein et al., 2008). Moreover, direct targeting of CDK8 with siRNA was shown to prevent the proliferation of colon cancer cells and to arrest cell cycle progression, additionally in a possible context with β -catenin down-regulation (He et al., 2011). Besides this cell-culture based evidence, CDK8 was shown to be a marker of poor prognosis for colon cancer patients (Firestein et al., 2010). Within its function as potent oncogene in colon cancerogenesis, CDK8 seems to act both Mediator-dependent and independent (reviewed by Firestein et al., 2009). On top of that CDK8 was as well linked to epigenetics (Kapoor et al., 2010). Only recently CDK8 was reported to contribute to maintaining both tumors and embryonic stem cells in an undifferentiated state. Thus targeting CDK8 therapeutically was suggested to specifically inhibit stem-cell like properties of cancer cells (Adler et al., 2012). Taken together these findings in literature classify CDK8 as an attractive target for drug development.

1.6 Drug-ability of CDK8/CycC: type I and type II inhibitors

Because of their important role in physiology and disease, extensive efforts towards inhibitors against CDK members have been performed. In a first approach compounds were developed to mimic and compete with ATP binding to a kinase (termed “type I” inhibitors). Such type I inhibitors (Figure 5) normally bind a kinase in an active state with an open conformation of the T-loop. This conformation

is normally characterized by the position of the conserved triad - aspartate - phenylalanine - glycine (DFG) - at the beginning of the activation segment with the phenylalanine side-chain rotated within the kinase binding site and the side-chain of the aspartate solvent-exposed (termed “DFG in”). Thereby the type I binding compounds mimic the binding mode of the adenine ring of the ATP to the region connecting the N-lobe with the C-lobe (termed “hinge”) (Zuccotto et al., 2010). However, within the CDK family the topology in the ATP-binding site is very well conserved generating promiscuity of type I inhibitors (Echalier et al., 2010). However, over the past years a new strategy for drug discovery has been developed caused by the clinical success of the drugs Nexavar[®] (sorafenib, BAY-43006, Bayer Pharma) and Gleevec[®] (imatinib, STI-571, Novartis Pharma AG). The efficiency of these drugs has been attributed to a type I differing binding mode targeting a pocket adjacent to the ATP binding site, termed the “deep pocket” (Dietrich et al., 2010, Backes et al., 2008). The “deep pocket” is accessible in protein kinases by the rearrangement of the DFG-motif from the active state (DFG-in) to the inactive state (DFG-out). Besides binding the ATP-binding site such as the type I inhibitors, those “type II” inhibitors (Figure 5) extend into the deep pocket (Backes et al., 2008, Zuccotto et al., 2010). In contrast to type I inhibitors that bind an activated kinase, the application of inhibitors with a deep pocket binding mode is meant to target the inactive state of a kinase (Backes et al., 2008). Typically these type II compounds compromise so called slow binding kinetics (Kroe et al., 2003) with an elongated “residence time” (Copeland et al., 2006). It was suggested that these allosteric ‘type II’ inhibitors provide enhanced potency *in-vivo* by their ‘binding kinetic signature’ (Müller et al., 2010) and duration of pharmacological target interaction. In fact residence time, defining the time by which an enzyme is occupied by a compound, is under discussion to exhibit a similarly important effect as its apparent affinity regarding half-inhibitory concentration IC_{50} or the dissociation constant (K_d) (Copeland et al., 2006, Tummino et al., 2008). Additionally, the improvement of residence time is reported to be a mean to improve the correlation of *in-vitro* and *in-vivo* efficiency of drugs (Lu et al., 2010).

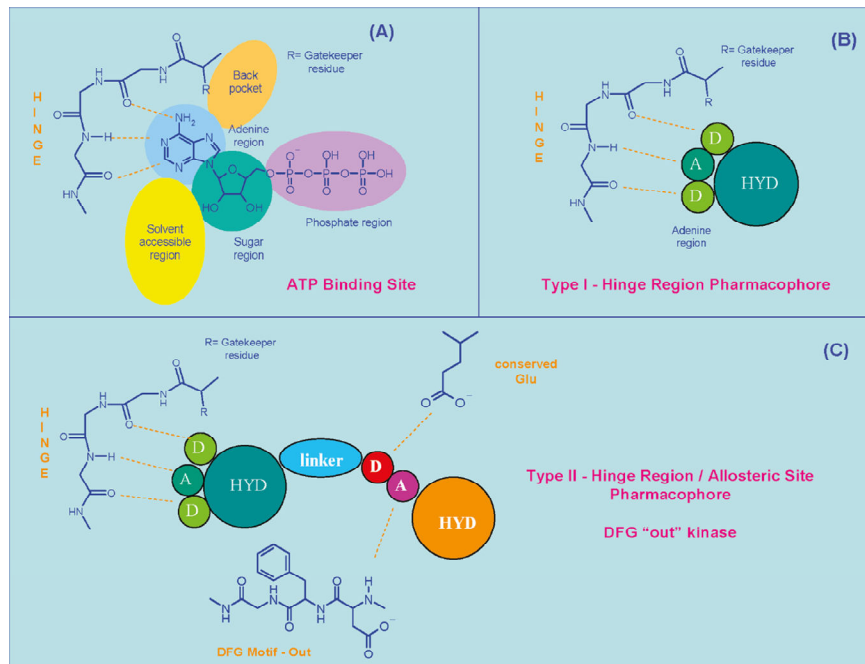


Figure 5 Binding mode of type I and type II inhibitors

Figure adapted from Zuccotto et al., 2010; D = H-bond donor, A = H-bond acceptor, HYD = hydrophobic moiety

Over the last decades the CDK/cyclin family was not believed to undergo the morphological changes in the ATP binding site that enable access to the adjacent deep-pocket for small-molecule binding (Betzi et al., 2011). Nevertheless, the deep pocket binder sorafenib (Nexavar[®], BAY-43-9006) (Wilhelm et al., 2006, Wan et al., 2004, Simard et al., 2009) was shown to bind CDK8/CycC (Neumann et al., 2011). This offered a new opportunity of targeting CDK8/CycC in its function as transcriptional regulator and potent oncogene.

1.7 Aims of this thesis

When this work was initiated, the CDK9/CycT complex (Baumli et al., 2008) was solved as the first transcriptional CDK/cyclin complex besides the previous structures of isolated CDK7 or isolated transcriptional cyclins such as the CycC yeast analogue SRB11. The transcriptional CDK9/CycT1 pair had already revealed important differences in its binding mode when compared to the cell cycle CDK/cyclin complexes. As the crystal structure of human CDK8/CycC complex was unknown at the time being, one main objective of this work was to determine the crystal structure of the human CDK8/CycC complex. Detailed structural information on the CDK8/Cyclin C heterodimer would provide further insight in differences between the transcriptional and cell cycle CDK/cyclin complexes. The crystal structure of the CDK8/CycC should provide information on how CDK8 gains CycC specificity in contrast to the other binding partner CDK3, especially because CDK/cyclin formation *in-vitro* is highly promiscuous. Furthermore the general two step model of the CDK activation should be investigated for CDK8/CycC as no phosphorylation of the CDK8 activation segment had been observed and it was important to find structural clues for a phosphorylation-independent activation of the CDK8 kinase.

The second objective of this work was to give structural evidence for the slow binding kinetics and long residence time of sorafenib towards CDK8/CycC (Neumann et al., 2011), to clarify whether sorafenib performs a deep pocket binding mode when complexed to CDK8/CycC. This would be quite surprising as no small-molecule induced DFG-out conformation of a CDK/cyclin complex had been observed. In any case, similar compounds with beneficial slow binding kinetics and an extended residence time should be developed in parallel to the respective crystal structures in a structure-kinetic-relationship (SKR) study since it is more and more accepted that a long residence time can be an important success factor for a drug. However, little is known about the interplay between the characteristics of the compound target interaction and binding kinetics. Such inhibitors could be a useful starting point to specifically inhibit CDK8/CycC in its role as important player in physiology and disease.

2. Material and Methods

2.1 Molecular Cloning

2.1.1 Polymerase chain reaction (PCR) based DNA amplification and sub-cloning

For standard polymerase chain reactions, a 50 μ l reaction with 500nM of forward and reverse primer, 500 μ M dNTP's (Roth), 1U Pfu Ultra II polymerase (Stratagene), 1x Pfu Ultra II buffer (Stratagene) and 10–50ng template DNA was set up. Amplification was performed in an Eppendorf MasterCycler gradient using the protocol given in Table 1.

95°C	3min	
95°C	20s	30 cycles
55-65°C	20s	
68°C	15s/kb	
68°C	3min	
4°C	∞	

Table 1 PCR program for standard DNA amplification

The PCR product was then purified using the QIAquick[®] PCR purification kit (Qiagen). After restriction enzyme digestion using New England Biolabs enzymes and protocols, the fragments were separated and purified from agarose gels with the QIAquick[®] gel extraction kit (Qiagen). The target vector was digested in the same way and also purified via gel-extraction. Ligation of the PCR product into the vector was carried out using 20ng of target vector and a three-fold molar excess of insert using T4 DNA Ligase (Fermentas) according to the manufacturers protocol. After transformation into *Escherichia coli* (*E. coli*) TOP10F' and growth at 37 C over night, plasmid DNA was prepared using the QIAcube (Qiagen) with the corresponding QIAprep[®] Spin Miniprep Kit (Qiagen). Orientation and integrity of the insert was confirmed by restriction enzyme digestion and DNA sequencing. For agarose electrophoresis and visualization of DNA typically, 0.5% to 2% (w/v) agarose were dissolved by heating in 50x TAE buffer (Applichem) and gels were prepared using a Sub-Cell GT electrophoresis system (Bio-Rad). After mixing with 6x DNA loading dye (Fermentas), the samples were loaded and electrophoresis was carried out at a constant voltage of 110V in 50x TAE buffer

(Applichem). Either a 1kb or 100bp DNA ladder (peqlab) was used as a size reference. DNA was visualized with SYBR[®] Safe DNA gel stain (Invitrogen), followed by imaging on a gel-doc system (Bio-Rad). CDK8 and CycC templates were obtained by GENEART for subsequent PCR. For protein expression in *E. coli* the CycC constructs including its full-length CycC(Hs1-283) and a C-terminally truncated mutant CycC(Hs1-261) were cloned into the pET28a vector (Novagen), that contains an N-terminal polyhistidine (HIS)-tag configuration including a Tobacco Etch Virus (TEV)- protease cleavage site. Other vectors used for protein expression of recombinant CycC-fusion protein in *E. coli* include the pGEX-6P-1 vector (GE Healthcare), containing an N-terminal glutathione S-transferase (GST)-fusion and a PreScission protease cleavage site or the pET32a vector (Novagen), that has an N-terminally thioredoxin (TRX)-fusion as well as a S-tag (STG)-fusion together with a thrombin (THB) cleavage site respectively an enterokinase (EK) cleavage site. For protein expression in insect cells C-terminally truncated construct variants (Hs1-464, Hs1-424, Hs1-403, Hs1-377 and Hs1-348) were evaluated besides the CDK8 full-length construct CDK8(Hs1-464) and its kinase-dead equivalent CDK8(Hs1-464 D173A). CDK8 constructs were cloned into the pFastBacHTa vector (Invitrogen) that had been previously modified including an insertion of a glutathione S-transferase (GST)-fusion with a PreScission protease cleavage site (gift by K. Maskos, Proteros Biostructures GmbH). The CycC full-length CycC(Hs1-283) and a C-terminally truncated mutant CycC(Hs1-261) was cloned into the pFastBacHTa vector with additional thioredoxin (TRX)-fusion and a S-tag fusion, including a thrombin cleavage site and an enterokinase (EK) cleavage site.

2.1.2 Bacterial transformation

Chemical transformation was used for all bacterial strains (Table 2) of this thesis. Transformation was essentially performed using standard protocols of the commercial distributors. Cells were thawed on ice and 10-50ng of DNA were added in a volume of 1-5 μ l. DNA was allowed to attach to the bacterial surface for 30min on ice and transformation was induced by heat-shock for 45s at 42°C. After 2 min incubation on ice, 300 μ l of SOC medium were added and bacteria were incubated at 37°C for 1 h to induce resistance. Cells were spread on plates carrying the respective antibiotic to select for transformants and incubated at 37°C over night.

Strain	Genotype	Application	Source
TOP10F'	F' { <i>lacIq Tn10</i> (TetR)} <i>mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) Φ 80 <i>lacZ</i> Δ M15 Δ <i>lacX74 recA1 araD139 Δ(<i>ara-leu</i>)7697 <i>galU galK rpsL endA1 nupG</i></i>	molecular cloning & DNA production	Invitrogen
DH10Bac TM	<i>E. coli</i> F ⁻ <i>mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) Φ 80 <i>lacZ</i> Δ M15 Δ <i>lacX74 recA1 endA1 araD139 Δ(<i>ara leu</i>)7697 <i>galU galK</i> λ- <i>rpsL nupG</i>/pMON14272/pMON7124</i>	molecular cloning & DNA production	Invitrogen
BL21-Rosetta TM (DE3)	<i>E. coli</i> F ⁻ <i>ompT hsdSB</i> (rB ⁻ mB ⁻) <i>gal dcm</i> (DE3) pRARE2 (CamR)	protein expression	Novagen
BL21-Codon Plus (DE3)-RIL	<i>E. coli</i> B F ⁻ <i>ompT hsdS</i> (rB ⁻ mB ⁻) <i>dcm</i> ⁺ Tetr <i>gal endA Hte</i> [argU ileY leuW Camr]	protein expression	Stratagene
BL21-pG-Tf2	<i>E. coli</i> BL21 F ⁻ , <i>ompT</i> , <i>hsdS_B</i> (rB ⁻ mB ⁻), <i>gal</i> , <i>dcm</i> with pACYC carrying <i>groES-groEL-tig</i> under Pzt1	protein expression	Takara
BL21-pTf16	<i>E. coli</i> BL21 F ⁻ , <i>ompT</i> , <i>hsdS_B</i> (rB ⁻ mB ⁻), <i>gal</i> , <i>dcm</i> with pACYC carrying <i>tig</i> under <i>araB</i>	protein expression	Takara

Table 2 Bacterial strains used in this thesis

2.2 Site-directed mutagenesis

Site-directed mutagenesis to insert a point mutation in order to generate the CDK8 kinase-dead mutant (D173A) was essentially performed as described in the Quick Change Site Directed Mutagenesis Kit (Stratagene). Primers were designed containing the desired mutation, flanked by 22 ± 2 nucleotides in each direction; preferentially ending in C/G. PCR based mutagenesis (Table 3) on the vector to be mutated was followed by a DpnI digest to eliminate the parental plasmid. After transformation into *E. coli* TOP10F' and plasmid preparation, clones were tested for successful mutagenesis by sequencing.

95°C	1 min	18 cycles
95°C	30 sec	
55°C	1 min	
68°C	15 min	
68°C	10 min	
4°C	∞	

Table 3 PCR for site-directed mutagenesis

2.3 Bacterial expression

Generally, a single colony of the corresponding *E. coli* strains was used to inoculate a 2ml pre-culture of LB medium (Roth) containing the respective antibiotic. For test-expression this colony was used to inoculate 100ml of LB medium respectively for 2l of LB-medium for up-scaled expression. Cells were grown in LB medium supplemented with the respective antibiotics at 37°C to an OD_{600nm} of 0.5-0.6, and then the temperature was lowered at 24°C respectively 19°C before expression was induced with IPTG. After growth for 4h at 24°C or 19h at 18°C, respectively, the cells were harvested by centrifugation for 10 min at 6000g in a Beckmann Coulter Avanti J-26XP centrifuge using a JLA-8100 rotor.

2.4 Insect cell culture and heterologous protein expression

For the expression of target protein in insect cells, a baculoviral stock, the viral stocks and the SF9 insect cells used for propagation of the virus were essentially handled according to the instructions of the Bac-to-Bac Expression System (Invitrogen). Thereby baculoviral stock was generated by transformation of DH10-BacTM cells with the corresponding CDK8-plasmids respectively CycC-plasmids for subsequent transposon mutagenesis. After blue/white selection free Bacmid DNA was isolated. During generation of the viral stocks (Figure 6) and protein expression, SF9 cells were propagated in SF900 serum free medium (Invitrogen). For expression of target proteins SF9 cells were titerless infected with baculoviral stock at a 1:10 dilution. During this process, SF9 cells were cultured in suspension using wave bags. After incubation for 72 h at 27 °C, cells were harvested by centrifugation.

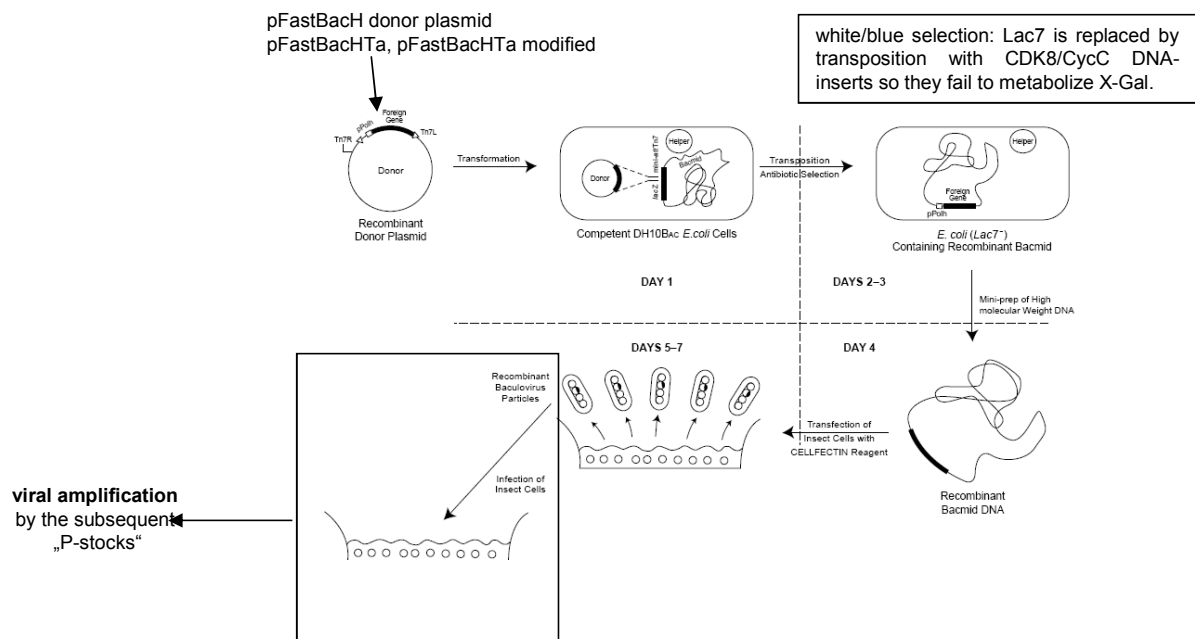


Figure 6 Generation of virus using the BAC-to-BAC system

Figure adapted from Invitrogen

2.5 Analysis of testexpressions

Analysis of test-expression in both the bacterial or insect cell system was analyzed using the NiNTa spin column kit (Qiagen) for HIS-tagged or TRX-HIS-STG-STG-EK-tagged target proteins respectively the GST SpinTrapTM prepacked microspin column kit (GE Healthcare). Test-purification was essentially performed as described in the instruction manual of the corresponding commercial distributor.

2.6 Sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE) analysis and protein staining

To denature protein samples, they were boiled in 1% sodium dodecyl sulfate (SDS)-loading buffer for 5min at 95°C. Subsequently samples were loaded on gels purchased from Invitrogen (NuPAGE[®] Novex[®] 10% Bis-Tris Midi Gel). Electrophoresis was performed at 190V until the running front reached the boarder of the gel. Unstained protein-marker II (peqlab) was used as a size reference

covering 10-200kDa. The gel running buffer was either 1x NuPAGE[®] MES SDS running buffer or 1x NuPAGE[®] MOPS SDS running buffer (Invitrogen). To visualize proteins, the SDS-PAGE gels were soaked in the Coomassie based staining solution InstantBlue[™] (Biozol) at room temperature (RT) for 30min. InstantBlue[™] protein-staining solution then was removed and substituted with water. Imaging was performed on a gel-doc system (Bio-Rad).

2.7 Western Blot Analysis

After SDS-PAGE analysis as described in the previous paragraph, Western Blotting was performed using the iBlot[™] Dry Blotting System (Invitrogen) with the proper iBlot[™] Gel Transfer Stocks PVDF regular (Invitrogen) basically according to the manual provided by the distributor. The anti-GST-peroxidase (Sigma Aldrich) was used as anti-body and visualized using PM BLUE POD Substrate, soluble (Roche).

2.8 Determination of Protein and Nucleic Acid Concentration

Protein and DNA coefficient concentrations were determined using a NanoDrop Spectrophotometer (ND-1000 spectrophotometer, peqlab). In order to calculate the concentrations of the protein samples of the different constructs (CDK8 and CycC variants, respectively, the CDK8/CycC complex variants) the respective extinction coefficients and theoretical molecular weights (MW) were calculated from primary proteins sequences using the ProtParam online tool (Gasteiger et al., 2005).

2.9 HIS-tag/GST-tag purification

Bacterial pellets or insect cell pellets were re-suspended in the respectively optimized lysis buffer for best yield and purity. Cells were homogenized and insoluble material was removed by centrifugation. The supernatant was purified using either the batch method (Ni Sepharose 6 Fast Flow[™] for HIS-tag or Glutathione Sepharose 4 Fast Flow[™] for GST-tag) or loaded onto either a 5 ml NiNTA column

(Qiagen) or a 5ml GStrap FF column (GE Healthcare) using the ÄKTA purifier system (GE Healthcare) with the purification step essentially performed as described in the instruction manual of the respective columns. Target proteins were eluted with the corresponding competitive reagent.

2.10 Size exclusion chromatography (SEC)

Proteins that were either purified by HIS-tag purification or GST-tag chromatography were concentrated and loaded onto a Hiloal 26/60 Superdex 200TM prep grade column (GE Healthcare) to purify the target protein by size subsequently to affinity chromatography. Thereby the ÄKTA purifier system was used with the purification step essentially performed as described in the instruction manual of the commercial distributor. A maximum sample size of 10ml was achieved by concentration of the samples. To estimate the MW of proteins and complexes, the columns were calibrated with the Low and the High Molecular Weight Gel Filtration Calibration Kit (Amersham Pharmacia Biotech), which use proteins in the range between 13.7 kDa and 669 kDa.

2.11 Proteins concentration steps

In order to concentrate protein samples after intermediate and final purification steps centrifugal filter devices (Amicon[®] Ultra, Millipore or Sartorius Vivaspin 500) with 10kDa, 30kDa or 50kDa nominal molecular weight limit were used as described in the provided protocol. Concentration was conducted up to the desired volume for intermediate purification and up to the desired protein concentration in final concentration steps.

2.12 Purification Protocols

All chromatography buffers were prepared and the pH adjusted at room temperature. Subsequently chromatography buffers were filtered (0.22 µm), degassed and cooled to 4°C prior to use. All purification steps were performed at 4°C on chromatography stations and columns obtained from GE Healthcare (ÄKTA system). Between purification steps, the protein was kept on ice or in a cold room/fridge at 4°C. Before the protein was pooled, samples were analyzed on pre-cast SDS-gels (10%) obtained from Invitrogen.

More than 12 different CDK8/CycC constructs were evaluated. In general all purification steps were performed at 4°C. Cells were harvested and disrupted in buffer A consisting of 20mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)/NaOH pH6.8, 500mM NaCl, 10% glycerol, 1mM tris(2-carboxyethyl)phosphine (TCEP), additionally containing Complete protease inhibitor cocktail (Roche). Lysate of recombinant CDK8 protein was purified using glutathione sepharose (GE Healthcare) with buffer A supplemented with 50mM reduced glutathione (GSH) for elution. Cleavage of glutathione fusions was performed over night with PreScission protease (Amersham Biosciences). Subsequently cleaved protein was recovered by negative affinity chromatography in buffer A and further using SEC with Superdex 200 26/60 (GE Healthcare). CycC was purified in buffer A using affinity chromatography (HIS-Trap FF, GE Healthcare) with subsequent removal of TRX-fusion by thrombin protease overnight. Protein recovery was performed using negative affinity chromatography followed by SEC. For *in-vitro* complex formation, both proteins were co-incubated for 1h, then the CDK8/CycC complex was purified using SEC.

2.13 Crystallization, X-ray data collection, processing and refinement

Crystal setups were performed either manually with the hanging drop set-up or with the crystallization robot (Phoenix) as sitting drops in 96-well plates at 20°C. Data were collected at the protein crystallography beam line X06SA at the Swiss Light Source (SLS), Villigen/Aargau. A rotation of 0.5° per frame and the maximum reasonable crystal to detector distance was chosen for data collection.

The first structure solution of the CDK8/CycC was achieved by data processing using the program MOSLFM (Leslie, 1992). Molecular replacement (MR) was performed using the automatic molecular replacement pipeline BALBES (Long et al., 2008). BALBES (Figure 7) is an online tool to which the CDK8/CycC target protein sequences as well as the MTZ-file were submitted manually. The BALBES pipeline uses a PDB-derived database that organizes the PDB-files by defining chains, domains and multimers of structures. This allows for model search with separately applied chains from different PDB-files (e.g. model consisting of domain1chain2 & domain2chain1). A system manager controls decisions on protocols according to search models. Programs used for MR and refinements include MOLREP, REFMAC and SFCHEK.

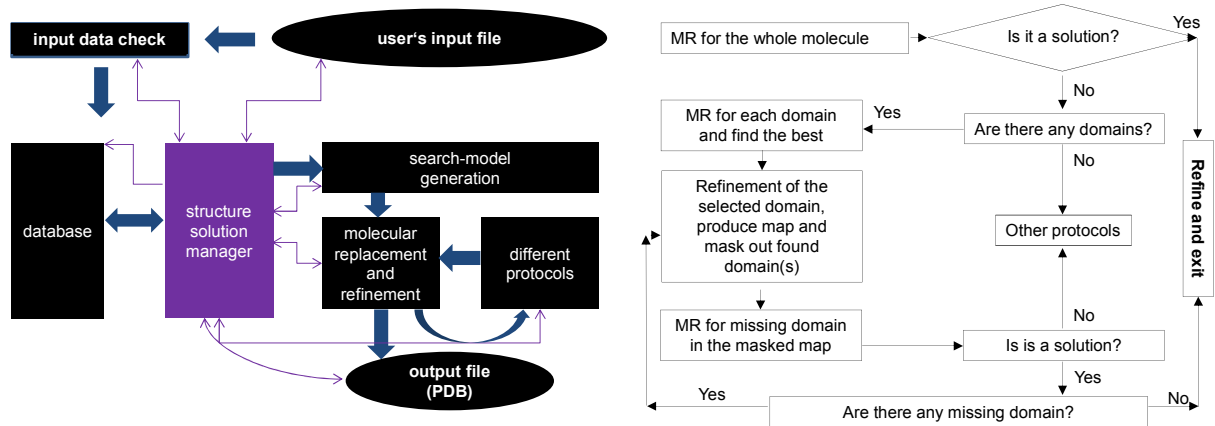


Figure 7 BALBES work-flow

Figure adapted from Long et al., 2008

Structure solution then was achieved by a combination of manual building and improvement of the electron density by the program Buccaneer (Cowtan, 2006). Buccaneer identifies connected α -carbon ($C\alpha$) positions using a likelihood-based density target in order to trace protein structures in a noisy electron density map. The same likelihood is applied in several ways to find candidate positions by a sixdimensional search, to grow a chain by adding new residues at either side of an existing $C\alpha$ position and to refine $C\alpha$ positions (see Figure 8). A more detailed description of how the phasing problem was solved for CDK8/CycC is found in chapter 3.5.

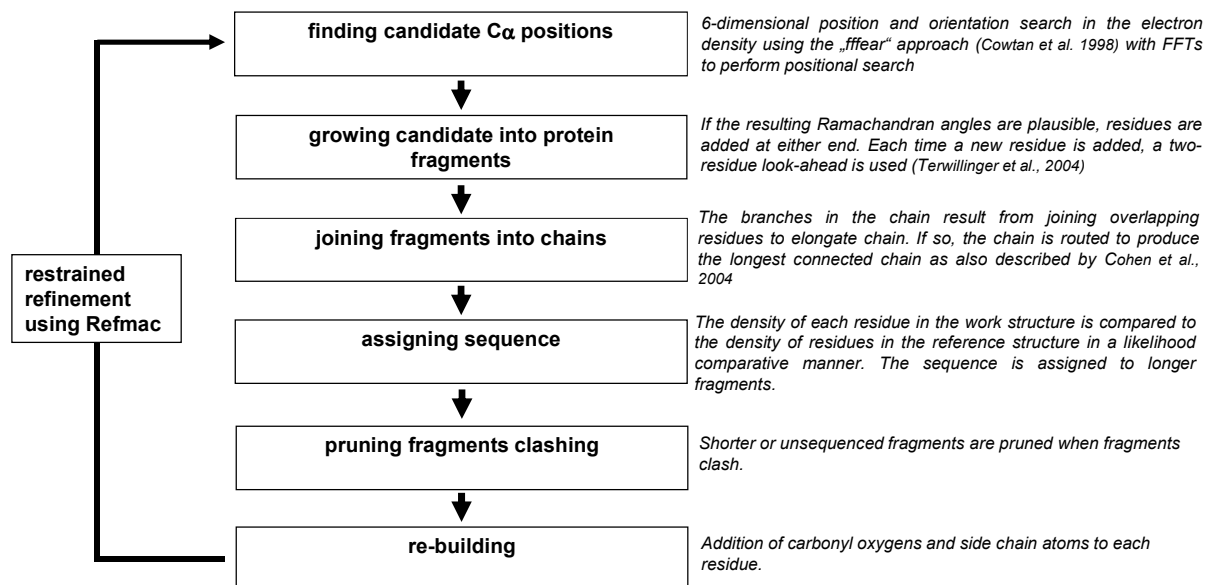


Figure 8 Work-Flow of Buccaneer

Figure adapted from Cowtan, 2006

The data of all other crystal complexes were processed and scaled using the program XDS (Kabsch, 1993) or processed using XDS and scaled using SCALA within the CCP4 program suite (Collaborative Computational Project 1994, Dodson et al., 1997). The phase information necessary to determine and analyse further structures of CDK8/CycC was obtained by molecular replacement with MOLREP (Vagin et al., 2010) using the first solved structure of CDK8/CycC (3RGF) as a search model.

Subsequent model building and refinement of all data sets was performed according to standard protocols with the software packages CCP4 and COOT (Emsley et al., 2004). For the calculation of the free R-factor, a measure to cross-validate the correctness of the final model, measured reflections were excluded from the refinement procedure. Translation liberation screw-rotation (TLS) refinement (using REFMAC5, Vagin et al., 2004) has been carried when thereby lower R-factors and higher quality of the electron density map was facilitated.

Solvent/water molecules were included in the model at stereochemically reasonable sites and all waters were checked with the validation tool of COOT. The location of the main-chain angles of non-glycine residues into the conformationally most favoured or allowed regions of the Ramachandran plot was checked by PROCHECK (Laskowski et al., 1993). Structures were overlaid using COOT/"SSM Superposition" (Krissinel et al., 2004). All sequence alignments were performed with ClustalW (Larkin et al., 2007). The secondary structure alignments were prepared using ESPript (Gouet et al., 1999), the graphical figures were made using PyMOL (The PyMOL Molecular Graphics System, Version 1.2, Schrödinger, LLC, New York, NY.). Calculation of the CDK8/CycC contact surface and identification of contact atoms of CDK8/CycC and a compound was performed with standard protocols using CCP4 structure analysis and corrected manually when stereo-chemically unreasonable.

2.14 CDK8/CycC K_d measurement

For affinity determination of the CDK8/CycC interaction, CDK8 was labelled with a fluorescent donor (340nm extinction/620nm emission), and CycC was labelled with a fluorescent acceptor (620nm extinction/665nm emission).

Binding between CDK8 and CycC was determined by measurement of the fluorescence resonance energy transfer (FRET) signal intensity at different CycC concentrations and of the binding constant calculated by non-linear regression, as fitted to Eq. (1):

$$\text{Fit} = ((BL_{\max} \times x) / (K_d + x)) \quad (\text{equation 1})$$

Where K_d is the dissociation constant, BL_{\max} is the emission at 665nm and x is the concentration of CycC.

2.15 The Proteros Reporter Displacement Assay

All experiments assessing the binding kinetics of the compounds were done with the Proteros Reporter Displacement Assay according to Neumann et al., 2009 and Neumann et al., 2011. In brief, the assay is based on the competitive displacement of a reporter probe designed to selectively target the CDK8 ATP-binding site with a fast binding kinetic signature. Binding of the probe to its target results in the emission of an optical signal. Competitive displacement of the probe by the corresponding compounds (fragment or lead-like type) results in a loss of the optical signal that can be quantified at increasing compound concentrations.

The IC_{50} measurement of the library screen was quantified by regular IC_{50} fitting of the percent reporter displacement values (with no displacement of the reporter corresponding to 0% (C^+) and complete displacement (C^-) corresponding to 100% after 15min and 80min). With the previously determined K_d value of the reporter and its known concentration, the K_d value of the compound is calculated within the equilibrated system from the measured IC_{50} -value by the Cheng-Prusoff equation (equation 2):

$$K_d(\text{compound}) = IC_{50} / (1 + [\text{reporter}] / K_d \text{ reporter}) \quad (\text{equation 2})$$

where $K_d(\text{compound})$ is the dissociation constant of the corresponding compound, IC_{50} is the concentration at which 50% of reporter binding are displaced, $[\text{reporter}]$ is the known concentration of the reporter probe and $K_d \text{ reporter}$ is the value of the previously determined dissociation constant of the reporter probe.

Binding kinetics was analyzed by fitting the reporter displacement traces for individual compound concentrations to a monoexponential decay equation. Thereby an exponent is yielded that reconciles the observed association rate k_{obs} for each concentration. Then each k_{obs} value is plotted on the y-axis against the corresponding compound concentrations (x-axis) and fitted to linear equation 3:

$$k_{\text{obs}} = k_{\text{off}} + k_{\text{on}} [\text{compound}] \quad (\text{equation 3})$$

where k_{obs} is the observed association rate constant, k_{off} is the dissociation rate constant, k_{on} is the association rate constant and [compound] is the compound concentration. Thus k_{on} is obtained from the slope.

The k_{off} -rate than is obtained by equation 3:

$$k_{\text{off}} = k_{\text{on}} \times K_{\text{d}} \quad (\text{equation 4})$$

where k_{off} is the dissociation-rate constant, k_{on} the association rate constant and K_{d} the dissociation constant.

The residence time is calculated from the dissociation rate in equation 5

$$\text{Residence time} = 1/k_{\text{off}} \quad (\text{equation 5})$$

Subsequently an example is given for compound X in Figure 9.

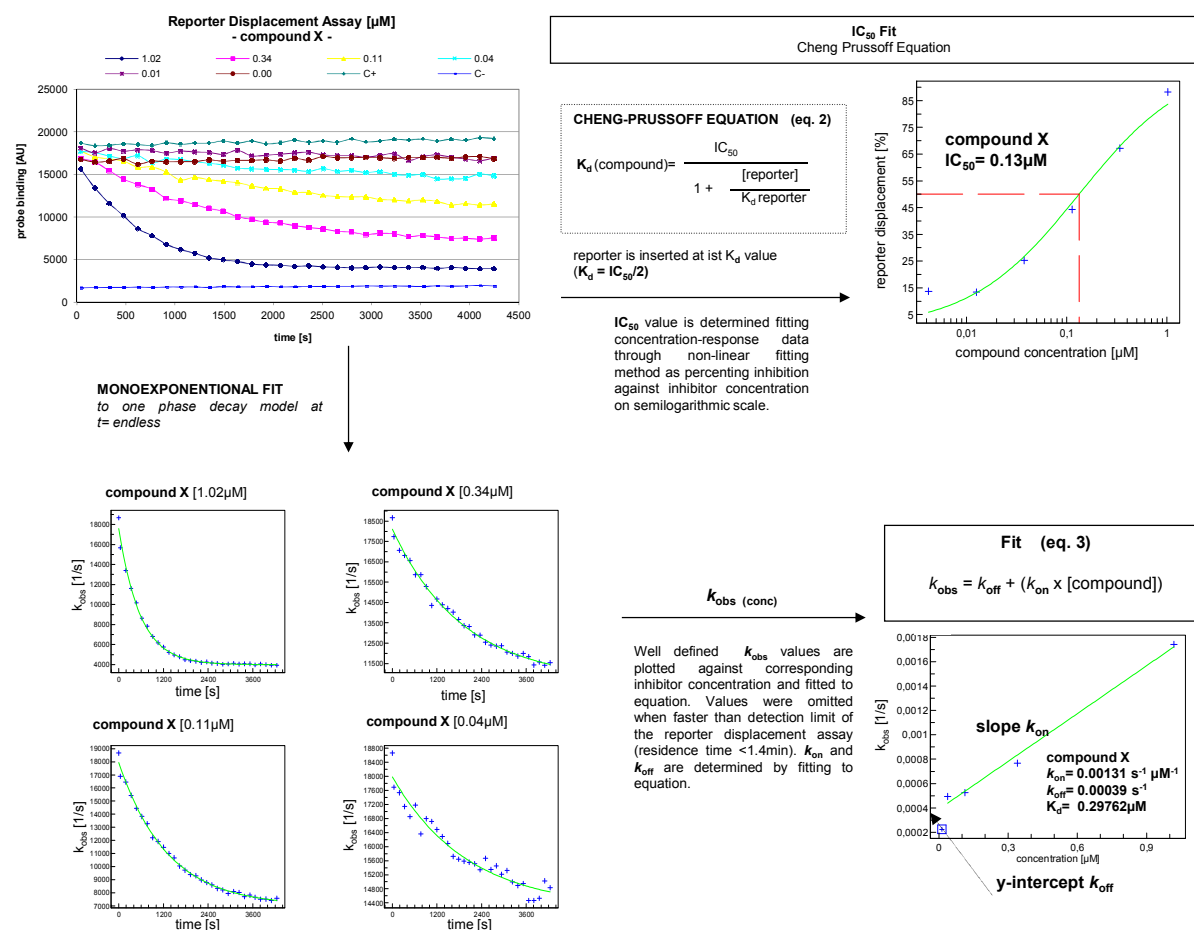


Figure 9 Schematic assay principle of the Reporter Displacement Assay

As reporter binding to its target results in the emission of an optical signal, the displacement of a competing compound can be quantified through the loss of the signal. The reporter displacement assay is applied to a compound X/ kinase Y interaction. Compound X is added at the desired compound concentrations [1.02, 0.34, 0.11, 0.04, 0.01 and 0.004 µM] and reporter displacement was measured continuously over time. No displacement of the reporter corresponds to 0% (C⁺) whereas complete displacement (C⁻) corresponds to 100%. The IC₅₀ value of compound X is quantified by regular IC₅₀ fitting using the percented reporter displacement values after the system has reached equilibrium. The K_d value is quantified by the Cheng-Prusoff equation (equation 2). Binding kinetics is determined by the application of a monoexponential decay function to the traces of the reporter displacement assay. Thereby an exponent is yielded that reconciles the observed association rate (k_{obs}) for the associated compound concentrations. These k_{obs} values are plotted against the corresponding compound X concentrations by fitting to equation 3. Thus k_{off} and k_{on} rates are obtained from the slope and the y-intercept. Residence time then is calculated by equation 5.

2.16 Chemicals

All materials were obtained from Sigma, Carl Roth or Applichem unless otherwise stated. Chemicals were >99% pure.

3. Establishment of protein production, crystallization and first structure determination of the human CDK8/CycC/sorafenib complex

3.1 Testing the different CycC-constructs in *E. coli* shows that it is not a suitable expression system for CycC but reveals the most appropriate tag-fusion construct

In a first step *E. coli* was tested as possible expression host for CycC starting with the GST-tagged CycC variants CycC(Hs1-283) and CycC(Hs1-261) cloned into pGEX-6P vector. Growth conditions were assessed at two different temperatures (24°C for 4h and 18°C for 19h). Within the first test-expressions, as exemplarily shown in Figure 10 the *E. coli* strains, BL21 Codon Plus RIL and BL21 Rosetta, were chosen that overexpress tRNAs rarely used in *E. coli*. However, the overall yield of CycC-fusion proteins was very low, suggesting the insolubility of the target protein at both temperatures (the overall similar effect is exemplarily shown in Figure 10). Moreover a contaminant was observable around 60kDa indicating complex formation of CycC with some *E. coli* host protein such as a chaperone. Subsequently a test-expression was performed in *E. coli* strains that co-express chaperones in order to clarify if this unspecific band at 60kDa (Figure 10) is caused by a chaperone-protein binding complex. Additionally the co-expression of chaperones facilitates proper protein folding and decreases inclusion body formation, possibly leading to increased solubility of the target protein. Besides the DnaK-DnaJ-GrpE and the GroEL-GroES folding modulator system Trigger Factor (Tf) is best characterized in *E. coli*. Tf binds the ribosomes and nascent polypeptide chains performing peptidyl-prolyl-cis/trans isomerase activity. It is under discussion whether Tf acts as general co-translational folding catalyst and stabilizer of nascent proteins before their transfer to chaperones. Furthermore the GroEL-GroES chaperonin system (ES and EL) belonging GroE chaperonins are mentioned to facilitate productive isomerization of proteins that already have reached a compact intermediate formation. Thereby substrates are bound and released in an ATP-dependent matter (Thomas et al., 1997). As the chaperone protein is reported to be bound to a transformed gene product, ATP and MgCl₂ can be added in order to separate this complex (Haga et al., 1998).

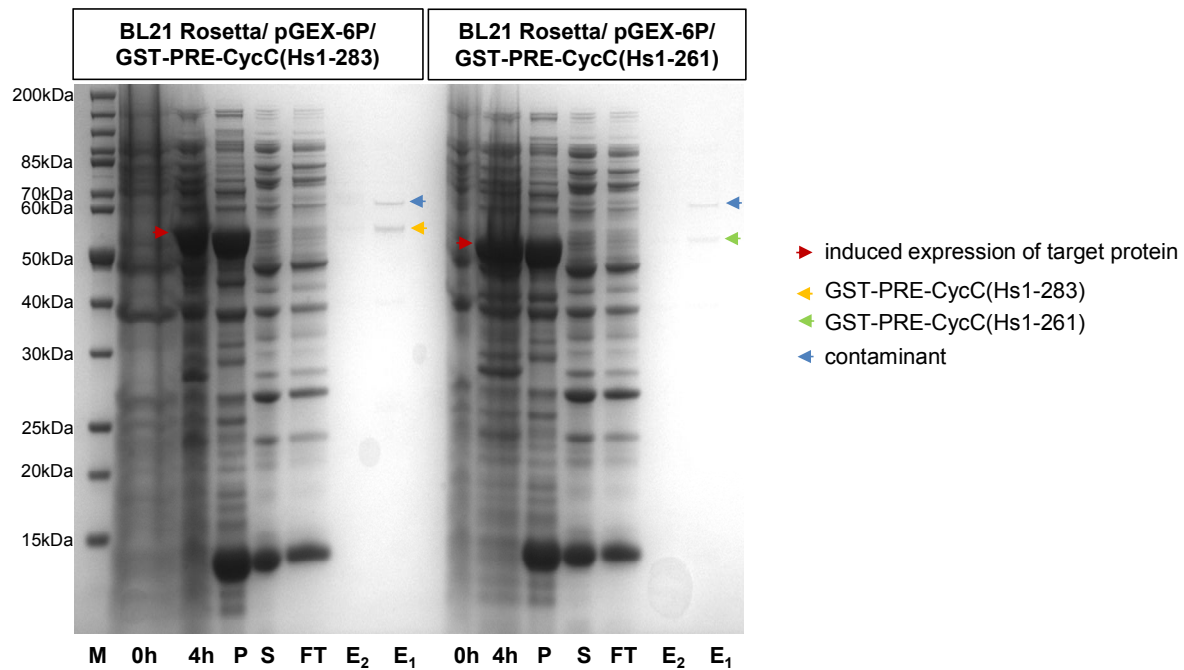


Figure 10 Example of the testexpression of GST-CycC-fusion protein in *E. coli* strains pRIL and Rosetta

10% SDS-PAGE analysis. The *E. coli* strain BL21 Rosetta was transformed with the CycC constructs (GST-PRE-CycC(Hs1-283) and GST-PRE-CycC(Hs1-261)). Protein expression was induced at 24°C for 4h with 1mM IPTG. Before induction (0h) and harvest (4h) a 1ml sample was taken and its loading on the gel normalized for absorption at OD_{600nm} . After harvest by centrifugation, cells were disrupted by adding phosphate buffered saline (PBS) additionally containing 1mM phenylmethylsulfonyl fluoride (PMSF) and 1.6mg/ml lysozyme, and ultrasonification. Cell lysate was then spun down to separate insoluble protein-fractions as pellet (P). Clarified supernatant (S) was used for purification by the GST SpinTrap™ prepacked microspin column kit (GE Healthcare). Cell pellet was resuspended in 6M urea buffer (P). After application of the supernatant to the GST-spin columns (flow-through was collected as FT), columns were washed twice with PBS. Elution was performed twice (E_1 , E_2) with 20mM reduced glutathione (GSH).

Even though the expression of CycC-fusion protein within chaperone-overexpressing bacterial strains (BL21 (DE3) Tf2 and Tf16) resulted in a remarkable increase of target protein expression (exemplified in Figure 11), the contaminant around 70kDa was still observable. It was obviously not possible to separate the target protein/ chaperone complex even while the co-incubation of ATP/MgCl₂.

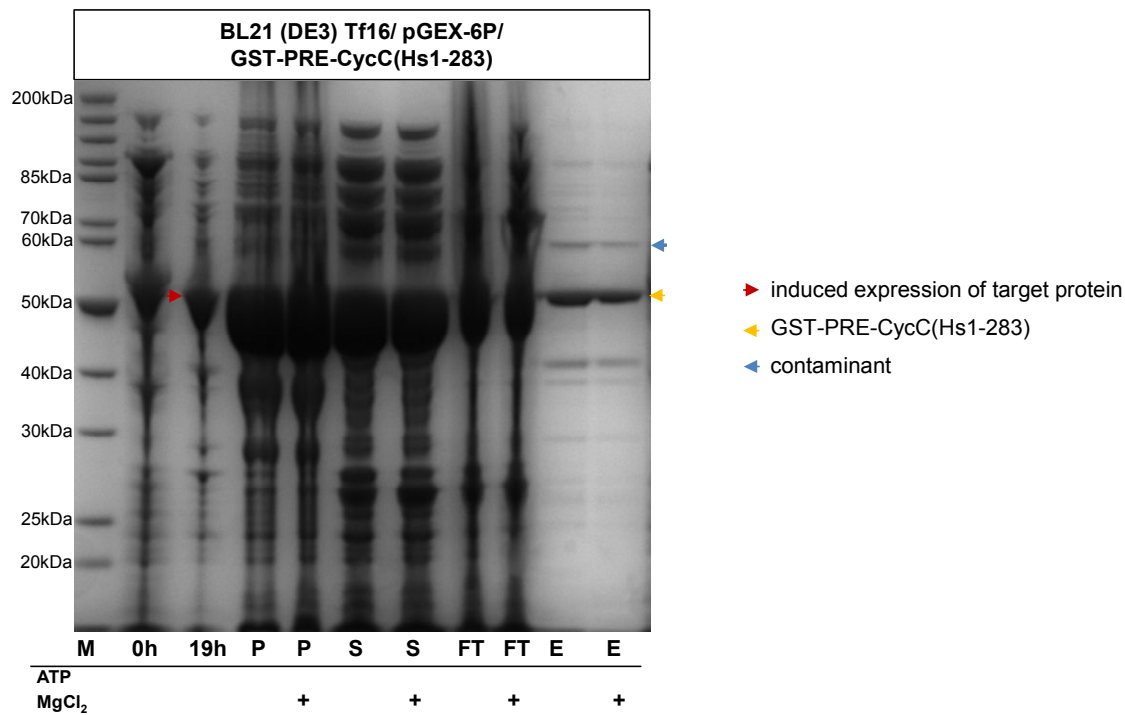


Figure 11 Example of the test expression of GST-tagged CycC fusion protein in *E. coli* strains that overexpress chaperones

10% SDS-PAGE analysis; *E. coli* strain BL21 (DE3) Tf16 was transformed with GST-PRE-CycC(Hs1-283). Immediately after inoculation chaperone expression was induced with 100mg/ml L-arabinose whereas protein expression of CycC was induced at 19°C for 21h with 0.1mM IPTG. Sample was taken before induction of protein expression (0h) and before harvest (4h). After harvest by centrifugation, cells were disrupted by adding PBS additionally containing 1mM PMSF and 1.6mg/ml lysozyme, and ultrasonification. Thereafter cell lysat were spun down and supernatant (S) was used for purification by GST SpinTrap™ prepacked microspin column kit (GE Healthcare). Cell pellet was resuspended in 6M urea buffer (P). After application of supernatant to GST-spin columns (flow-through was collected as FT), columns were washed twice with PBS. Elution was performed (E) with 20mM GSH. The purification was in parallel performed with the same buffers additionally containing 5mM ATP and 10mM MgCl₂ as indicated in the figure (+).

In summary, evaluation of GST-tagged CycC constructs did not result in a sufficient production of heterologous protein: the target proteins were insoluble, trapped in the protein pellet and seemed to be degraded even though the expression strains used were all protease-deficient. Lowering the temperature from 24°C to 18°C and over-expression of rare tRNAs (in expression systems BL21 (DE3) Codon Plus RIL/ Rosetta) was not efficient to increase solubility of target proteins. Co-expression of chaperones (in expression systems BL21 (DE3) Tf2/Tf16) resulted in a positive effect, but it was not possible to remove the recombinant protein from chaperone-contaminants by adding ATP/MgCl₂. Alternatively the maltose-binding protein tag, the N-utilization substance tag and the thioredoxin (TRX) - tag are reported to be most efficient in increasing solubility of heterologous

proteins in *E. coli* (Esposito et al., 2006). Therefore CycC-constructs (Hs1-283, Hs1-261) were cloned into the pET32A vector that includes an N-terminal thioredoxin (TRX) – tag in combination with a poly-histidine (HIS) - tag, a thrombin protease (THB) – cleavage-site, an S (STG) – tag and an enterokinase (EK)– cleavage site. Thereby the additional combination of the TRX-tag with a HIS-tag of the fusion-protein at the N-terminus allows for immobilized metal ion affinity chromatography purification (IMAC). The generation of the TRX-tag containing CycC-constructs showed a remarkably positive effect on the solubility of proteins and resulted in an extremely increased yield of protein production (Figure 12). Moreover, the contamination of eluted target protein by host proteins was clearly reduced.

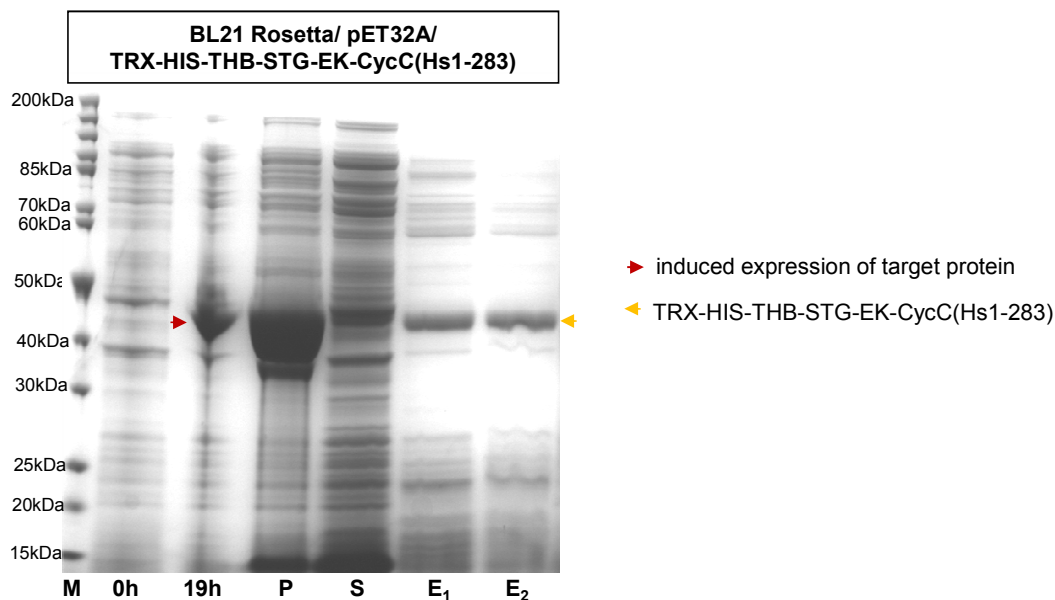


Figure 12 Test-expression of TRX-tagged CycC-constructs in BL21 (DE3) Rosetta

10% SDS-PAGE analysis; The *E. coli* strain BL21 Rosetta was transformed with TRX-HIS-THB-STG-EK-CycC(Hs1-283). Induction of protein expression was induced at 24°C for 4h with 1mM IPTG. Before induction (0h) and harvest (19h) a 1ml sample was taken and its loading on the gel normalized for absorption at OD_{600nm}. Cells were harvested by centrifugation and disrupted under non-denaturing conditions with ultrasonification in 50mM Na₂HPO₄/NaH₂PO₄, 300mM NaCl and 10mM imidazole (pH8.0) additionally containing 1mM PMSF and 1.6mg/ml lysozyme. Thereafter cell lysate were spun down and supernatant (S) was used for purification by was applied to NiNTa Spin column kit (Qiagen) whereas protein pellets were solubilized in 6M urea (P). Columns were washed twice with 50mM Na₂HPO₄/NaH₂PO₄, 300mM NaCl and 10mM imidazole (pH8.0). Elution was performed twice (E₁, E₂) with 50mM Na₂HPO₄/ NaH₂PO₄, 300mM NaCl and 250mM imidazole (pH8.0).

The fusion with the TRX-tag was most efficient compared to the GST-fusion proteins used in different *E. coli* strains before. This is congruent with a study of different fusion-proteins for the expression of different mammalian full-length proteins and protein domains that revealed that the TRX-tag is the best N-terminal fusion partner promoting soluble expression in *E. coli* besides the MBP-tag (Dyson et al., 2004). Subsequently, protein expression of TRX-HIS-THB-STG-EK-CycC(Hs1-283) and TRX-HIS-THB-STG-EK-CycC(Hs1-261) in BL21 Rosetta was up-scaled for purification (Figure 13).

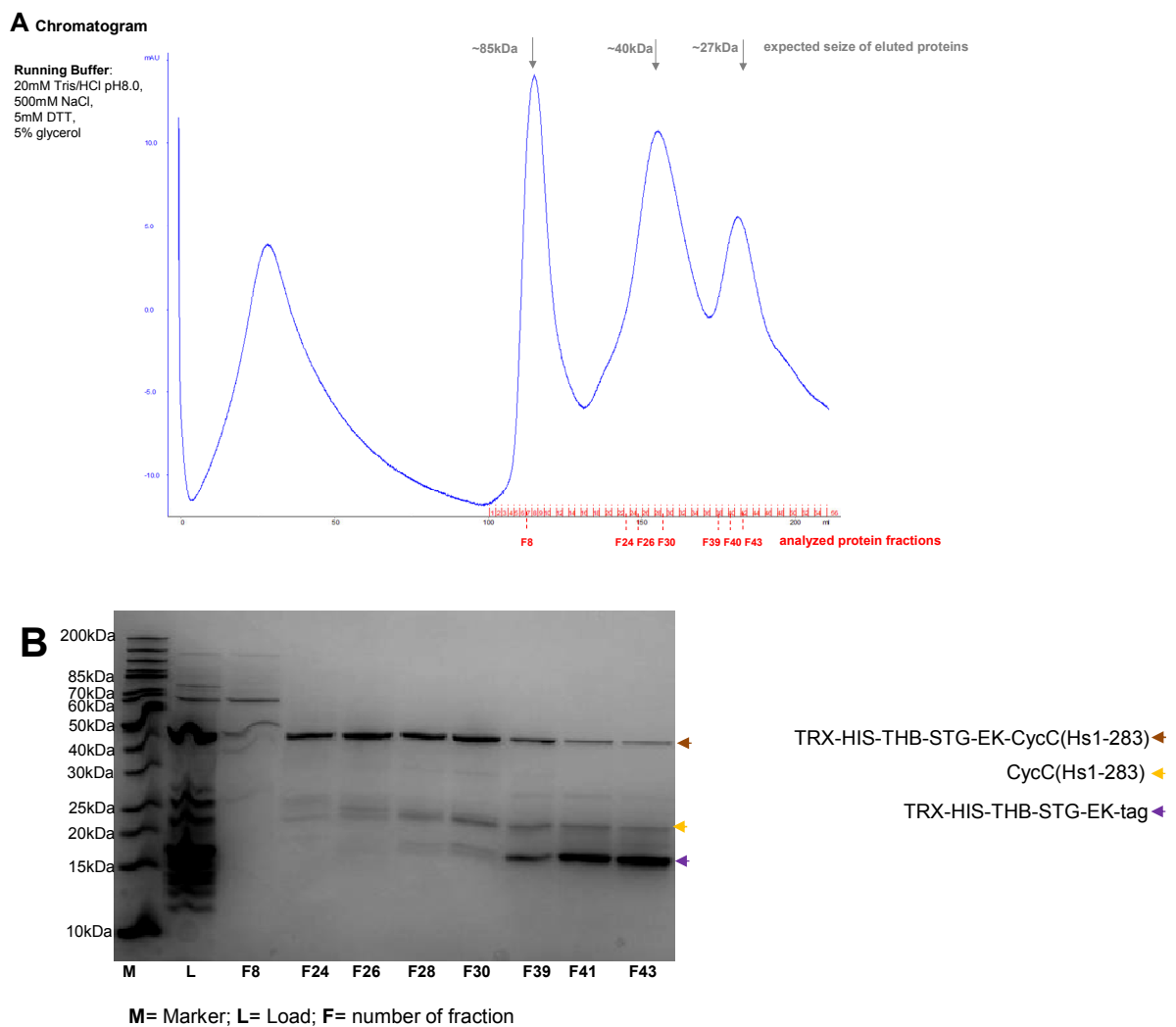


Figure 13 Purification of *E. coli* expressed CycC (TRX-HIS-THB-STG-EK-CycC(Hs1-283))

(A) Chromatogram of the SEC (Superdex 75 26/60, GE Healthcare) and (B) SDS-PAGE analysis (10% SDS) of the third step of CycC purification: analytical SEC after proteolytical removal of the TRX-tag; retentions volume corresponding seize is indicated with black arrows; protein fractions with red arrows and red labels; purified proteins fractions are marked on SDS-PAGE analysis with gold arrow (CycC(Hs1-283)), a brown arrow (TRX-HIS-THB-STG-EK-CycC(Hs1-283)) and a purple arrow (TRX-HIS-THB-STG-tag).

Even though a reasonable yield of target protein was achieved (9.5mg/ml in 20ml after positive affinity chromatography), the subsequent proteolytical digestion of the tag resulted in precipitation of the target protein as shown within the following analytical SEC and the corresponding PAGE-analysis (Figure 13). Thereby the higher part of purified protein is uncleaved TRX-CycC whereas there is only a small fraction of tag-removed CycC, which underlines that the removal of the tag leads to precipitation of CycC.

The precipitation of recombinant CycC target protein was blamed on the bacterial expression system that lacks posttranslational modifications such as phosphorylation or glycosylation and a sufficient folding machinery, resulting in insoluble or non-functional eukaryotic target proteins. In these cases overexpression in a eukaryotic system is advisable such as in the insect cell system. As a consequence a new vector was generated by modification of the vector pFastBacHTa with an N-terminal insertion of TRX-THB-STG-EK in order to facilitate CycC-fusion protein production in insect cells. Cloning was performed according to standard procedure and approved by sequencing; baculovirus production was successfully established as described in material and methods.

3.2 Protein expression of CDK8 and CycC is only possible separately in SF9 cells

As protein-expression in *E. coli* was not feasible, it was intended to express the CDK8/CycC target protein complex recombinantly in SF9 insect cells with the newly generated vector containing the TRX-THB-STG-EK-tag. Whereas it was possible to express isolated GST-tagged CDK8 and isolated HIS-tagged CycC, first trials to co-infect cells with virus of GST-tagged full-length CDK8(Hs1-464) and full-length CycC(Hs1-283) did not lead to any detectable protein production of the recombinant protein complex (similar effect is observed in Figure 14). Thus it was reasoned that the quality of the recombinant proteins is possibly affected by proteolytical activity of proteases that are often simultaneously expressed with the heterologous target proteins (Gotoh et al., 2001). Thereby proteases are reported to be produced during the whole period of target protein expression in insect cells: first as a stress response of the cells towards virus infection, second during the infection cycle by the baculovirus vector itself encoding cystein proteases and finally following cell lysis (Ikonomou et al., 2003). Cysteine protease activities are supposed to be predominantly in infected cell culture, followed

by those of aspartic proteases (Ikonomou et al., 2003). Thus subsequently inhibitors were tested to control proteolysis in baculovirus expression systems such as the SF9 cells (Martensen et al., 2001) and to increase stability of recombinant proteins expressed in insect cell system. These include the cysteine protease inhibitors E64 (N-[N-(L-3-trans-carboxyoxirane-2-carbonyl)-L-leucyl]-agmatin) during the protein expression and leupeptin as well as the proteasome inhibitor MG132 for purification. There was no positive effect on the protein expression of CDK8/CycC by the supplementation of the inhibitors (Figure 14).

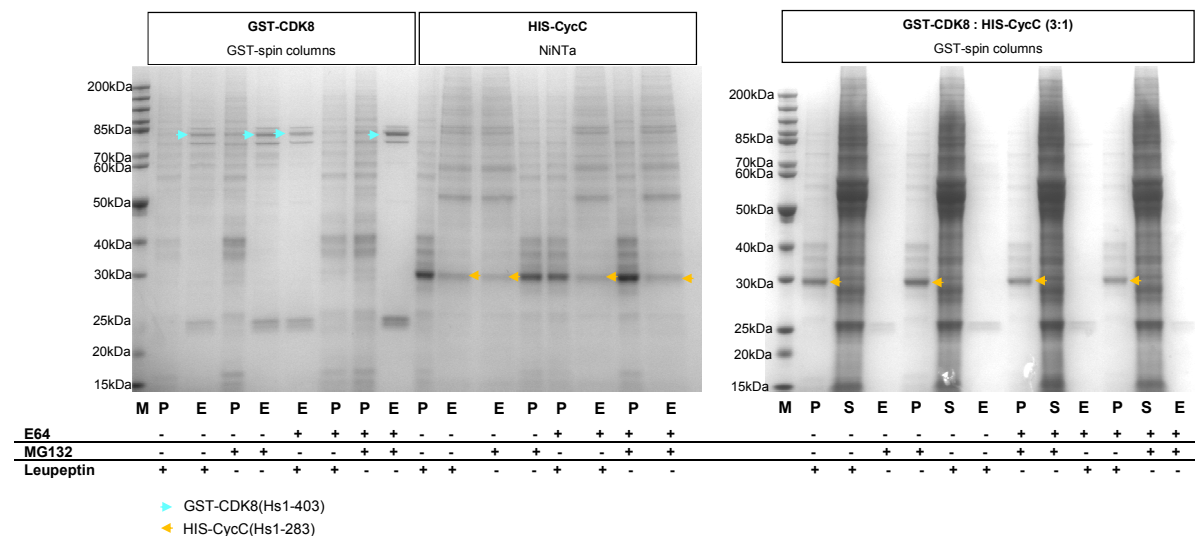


Figure 14 Test-expression of CDK8/CycC in SF9 cells in presence of proteolytic inhibitors

M= marker, P= pellet, E= eluate, S= supernatant; GST-CDK8= GST-PRE-CDK8(Hs1-464); HIS-CycC= HIS-TEV-CycC(Hs1-283); equivalent amounts of baculovirus to SF-9 cells subconfluently grown in 75cm² flasks. After 60h cells were harvested and disrupted by centrifugation. Supernatant was used for purification using GST spin columns (Qiagen) or NiNTA spin columns (Qiagen). Therefore cells were spun down, cleared supernatant (S) was applied to columns whereas pellet was resolved in 6M urea (P). After application of supernatant to GST-spin columns columns were washed twice with PBS. Elution was performed (E) with 20mM GSH. After application of supernatant to NiNTA-spin columns, columns were washed twice with 50mM Na₂HPO₄/NaH₂PO₄, 300mM NaCl and 10mM imidazole (pH8.0). Elution was performed twice (E1, E2) with Na₂HPO₄/NaH₂PO₄, 300mM NaCl and 250mM imidazole (pH8.0). Inhibitors (E64: 20µg/ml, MG132: 10µM, Leupeptin: 100µM) were added as marked (+). E64 was added to medium 24h after infection, MG132 and Leupeptin were added to lysis and analysis buffers.

However, while during co-infection the stability of GST-CDK8 was not improved by the presence of proteasomal inhibitors, a tendency of increased accumulation of insoluble CycC target protein was observable within the cell pellet (Figure 14). As CycC is degraded in an ubiquitylation-proteasome dependent manner (Barette et al., 2001), its accumulation seems quite plausible. According to these

results the mutual destabilizing effect of CDK8/CycC during protein production was not blamed to the proteolytic machinery in insect cells or proteasomal degradation. A negative influence of the CDK8/CycC on insect cell physiology seemed more likely.

As the expression of a possibly active CDK8/CycC complex seemed to have a cytotoxic effect on insect cell physiology, a kinase-dead mutant CDK8 (Hs1-464D173mutA) was assessed for co-infection with CycC. Mutating the CDK8 residue D173 to A173 was reported to abrogate its kinase activity and TFIID-dependent transcriptional repression (Akoulitchev et al., 2000). In addition to the previously assessed HIS-tagged CycC-constructs, the TRX-tagged CycC variants (TRX-HIS-THB-STG-CycC(Hs1-283) and TRX-HIS-THB-STG-CycC(Hs1-261)) were evaluated accordingly to the promising results of the test-expression of CycC-fusion protein in *E. coli* (see chapter 3.1). Thereby an increased stability and solubility of recombinant TRX-tag fused CycC protein had been revealed. For the *in-vivo* complex formation, both CDK8 full-length mutants, active (GST-PRE-CDK8(Hs1-464); referred to as GST-CDK8+) and inactive (GST-PRE-CDK8(Hs1-464 D173A); referred to as GST-CDK8 D173A) were tested in combination with different ratios of the HIS-tagged and TRX-tagged CycC baculovirus. The combination of the TRX-tag with an N-terminal HIS-tag allows for IMAC purification of the TRX-tagged CycC proteins.

As a result, for *in-vivo* complex formation with CycC SDS-PAGE analysis only detected a band at the CDK8 corresponding size for cells that were co-infected with the kinase-dead CDK8 mutant. Even though only the CycC-fusion protein contains a HIS-tag it was possible to additionally purify GST-CDK8 D173A using the NiNTA module which indicates the *in-vivo* complex formation between the GST-CDK8 D173A/ the CycC-constructs (Figure 15A). Thereby TRX-tagged CycC obviously does not negatively affect the CDK8/CycC complex formation although the MW of TRX is comparatively larger than that of the previously assessed HIS-tag. Subsequently western blot analysis clarified the identity of the GST-CDK8 D173A band (Figure 15B) using an antibody against the GST-CDK8 fusion proteins. Interestingly the CDK8 kinase dead mutation does not inhibit complex formation with CycC. This is consistent with results from literature suggesting that CDK8/CycC complex formation occurs independently of catalytic activity of CDK8 (Barette et al., 2001).

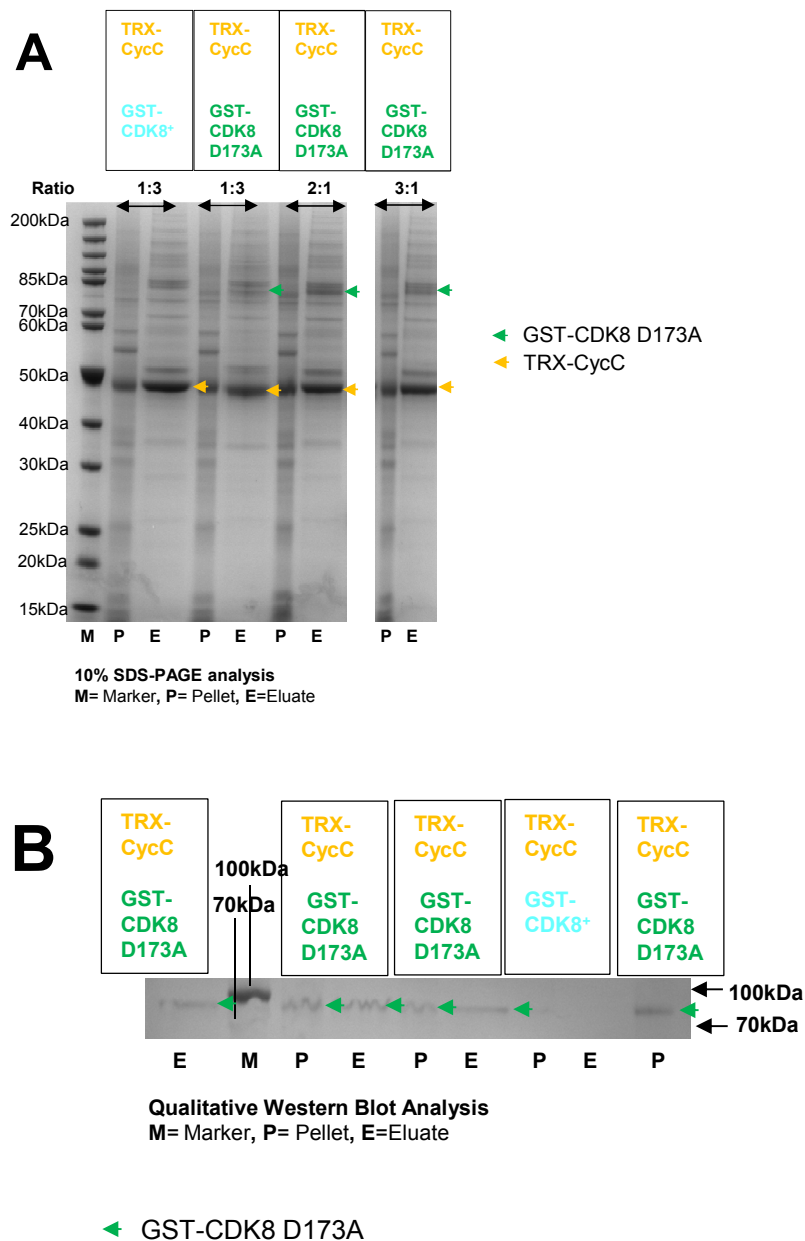


Figure 15 CDK8/CycC *in-vivo* complex formation using an active and a kinase-dead CDK8 mutant

GST-CDK8⁺ = GST-PRE-CDK8(Hs1-464); GST-CDK8 D173A= GST-PRE-CDK8(Hs1-464 D173A); TRX-CycC = TRX-HIS-THB-STG-EK-CycC(Hs1-283); equivalent amounts of baculovirus (as indicated by ratio) was applied to SF9 cells subconfluently grown in 75cm² flasks. After a period of 72h cells were harvested and disrupted by centrifugation. Supernatant was used for purification using NiNTA spin columns (Qiagen). After application of supernatant to NiNTA-spin columns, columns were washed twice with 50mM Na₂HPO₄/NaH₂PO₄, 300mM NaCl and 10mM imidazole (pH8.0). Elution was performed twice (E1, E2) with Na₂HPO₄/NaH₂PO₄, 300mM NaCl and 250mM imidazole (pH8.0). Samples were analyzed by 10% SDS-PAGE analysis (A) or proceeded according standard Western Blot analysis protocol (iBlotTM Invitrogen) using a GST-antibody.

In sum, the data of CDK8/CycC *in-vivo* complex formation suggest that the catalytically active CDK8/CycC complex cannot be achieved in the insect cell system. Despite of the difference of

species, expression of the human CDK8/ CycC complex (drosophila CDK8/ human CDK8 79.1% homology; drosophila CycC/ human CycC 72% homology) could disadvantageously influence the cell cycle and transcriptional process of the insect cell system. Even though there is no CDK8-analogue characterized for SF9 cells, the casein kinase II subunit alpha (UNP ID: O76484) was found to show at least some sequential identity with CDK8 (human CDK8: 25.6%; drosophila CDK8: 25.1%). Interestingly casein kinase II subunit alpha possibly participates in wnt signaling (UniProtKB), which is a functional link to human CDK8/CycC: together with MED12 and MED13, the CDK8/CycC mediator submodule is involved in wnt-signaling as well (Kim et al. 2006). Despite of the observation that *in-vivo* co-expression and complex formation of CycC with the kinase-dead CDK8 construct seems to be feasible, the yield of target protein complex obviously is not sufficient. As a consequence CDK8/CycC complex formation had to be achieved *in-vitro* with both target proteins being expressed isolated in SF9 cells.

3.3 Establishment of the CDK8/CycC complex formation *in-vitro* and its purification

During the establishment of the CDK8/CycC complex *in-vitro* and its purification, unexpected problems occurred. Even though GST-CDK8 constructs were already shown to be stable without tag when purified isolated or combined with HIS-TEV-CycC(Hs1-283), tag-cleaved CDK8 constructs precipitated when co-incubated with the TRX-HIS-CycC fusion protein for complex formation. It was not possible to re-solve the target protein complex by co-incubation with different detergents. Therefore it seemed necessary to additionally stabilize the target protein complex using the CDK8/CycC specific inhibitor sorafenib during the whole purification process. Optimization of the purification protocols resulted in a final purification protocol for CDK8/CycC (Figure 16). This protocol seemed suitable for all different CDK8 and CycC variants and for *in-vitro* complex formation of CDK8/CycC and resulted in target protein feasible for crystallization screening of the CDK8/CycC complex.

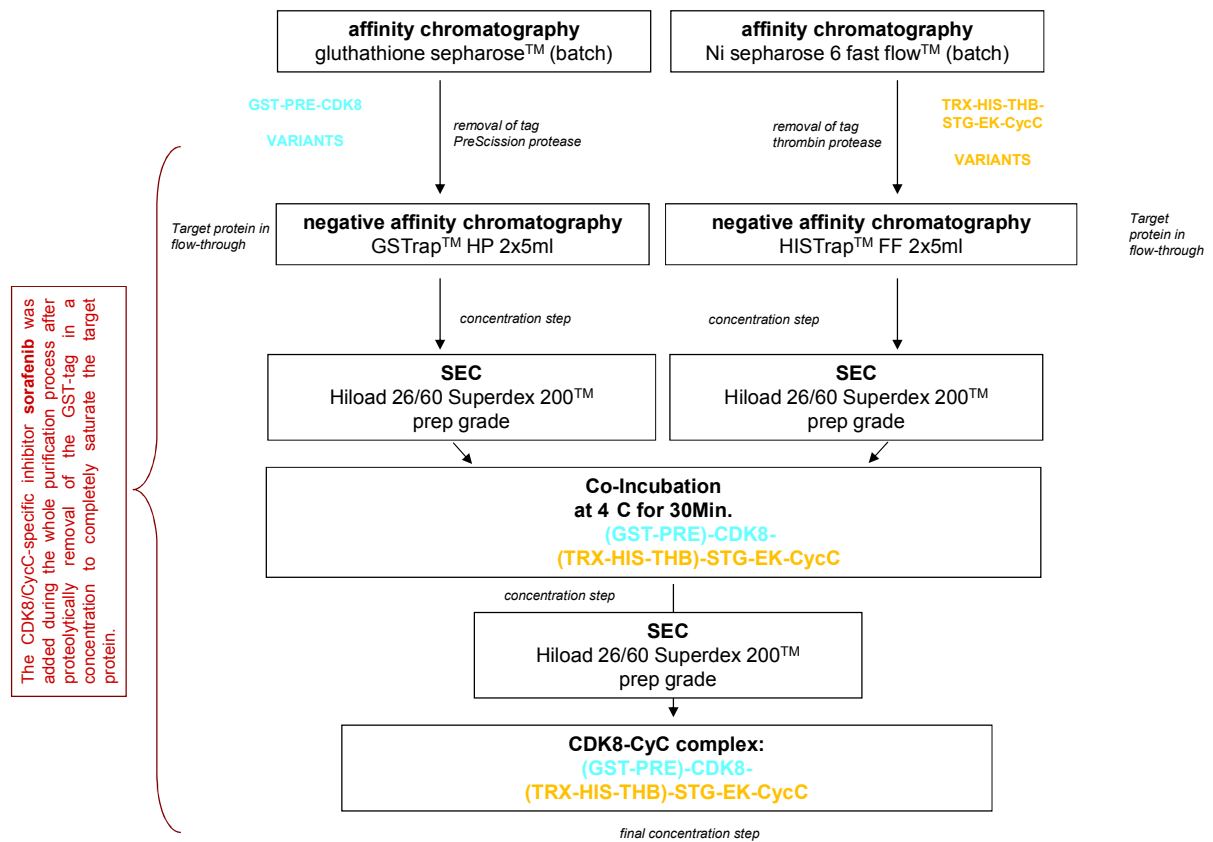


Figure 16 Final protocol for purification of the CDK8/CycC protein complex assessing different CDK8 and CycC constructs

In the following an example of successful *in-vitro* complex formation is given as justified by the final purification step using SEC. Thereby SEC revealed complex formation of CDK8(Hs1-403) with CycC(Hs1-283) which could be separated from CycC in excess and the GST-contaminant (Figure 17).

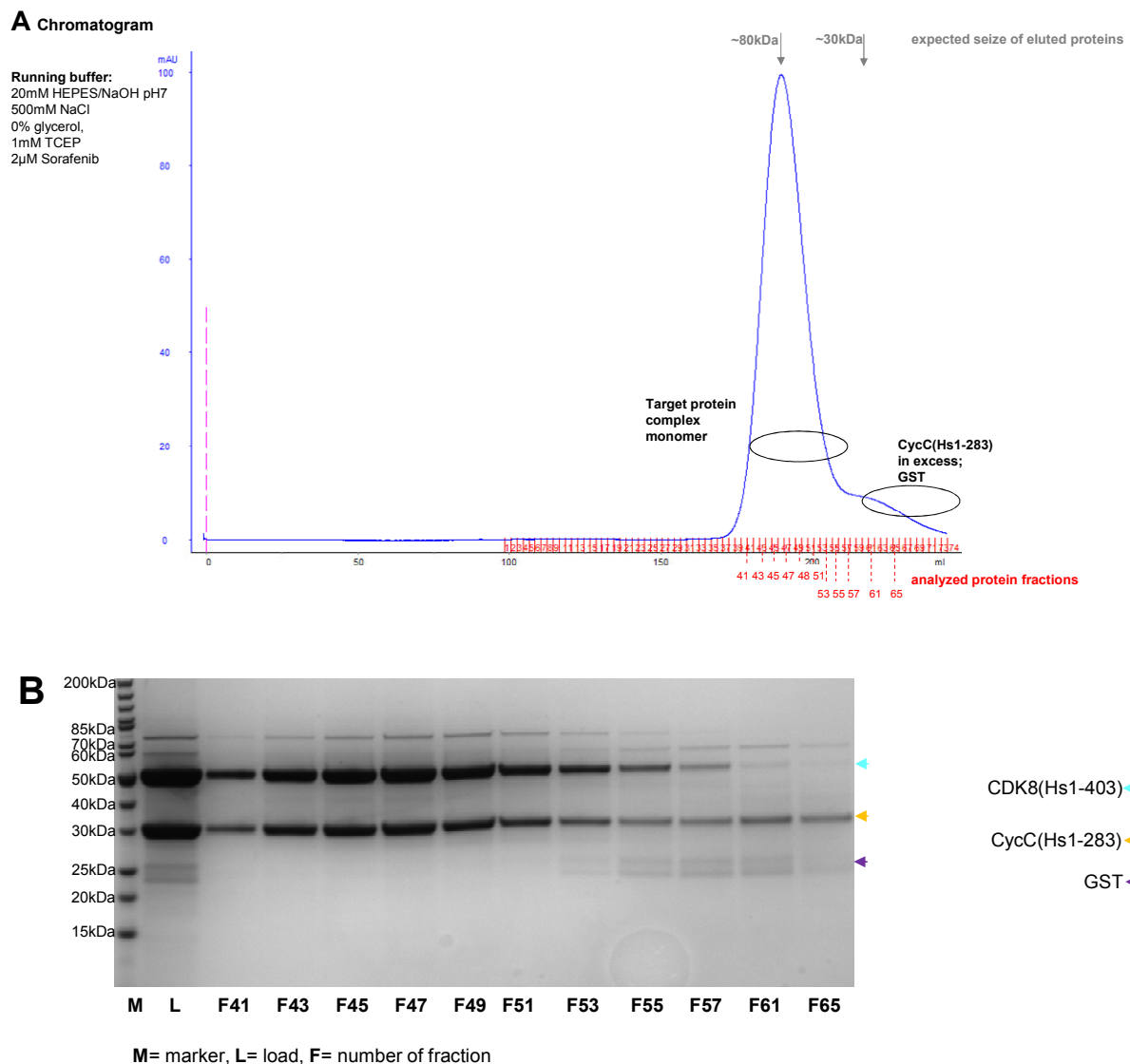


Figure 17 SEC purification of the in-vitro formed CDK8/CycC complex

(A) Chromatogram of the SEC (Hiload 26/60 Superdex 200TM prep grade column, GE Healthcare) and (B) SDS-PAGE analysis (10% SDS) of the final step of CDK8/CycC purification to purify the target protein complex; retentions volume corresponding seize is indicated with black arrows; protein fractions with red arrows and red labels; purified proteins fractions are marked on SDS-PAGE analysis with turquoise arrow (CDK8(Hs1-403), gold arrow (CycC(Hs1-283), and purple arrow (GST).

3.4 Crystallization and refinement of crystals of the CDK8/CycC/sorafenib complex

First diffracting protein-crystals were obtained from the constructs CDK8(Hs1-403)/CycC(Hs1-283) at a concentration of 11.3mg/ml, co-crystallized with 10μM sorafenib. Crystals were obtained after the addition of water to the reservoir (1:1) in the buffer containing 20% (w/v) PEG3350 and 0.2M LiCl at 20°C (Figure 18, left). In order to test whether crystallization had any impact on the *in-vitro* formed

CDK8/CycC complex, e.g. resulting in disintegration of target protein complex, SDS-PAGE analysis was performed. Therefore one crystal was isolated and dissolved in bi-distilled water before application within SDS-buffer on a gel. A sample of the same protein used for crystallization screening served as positive control (Figure 18, right). According to SDS-PAGE analysis, the *in-vitro* complex formation CDK8/CyC is not inhibited or reversed by the protein crystallization conditions.

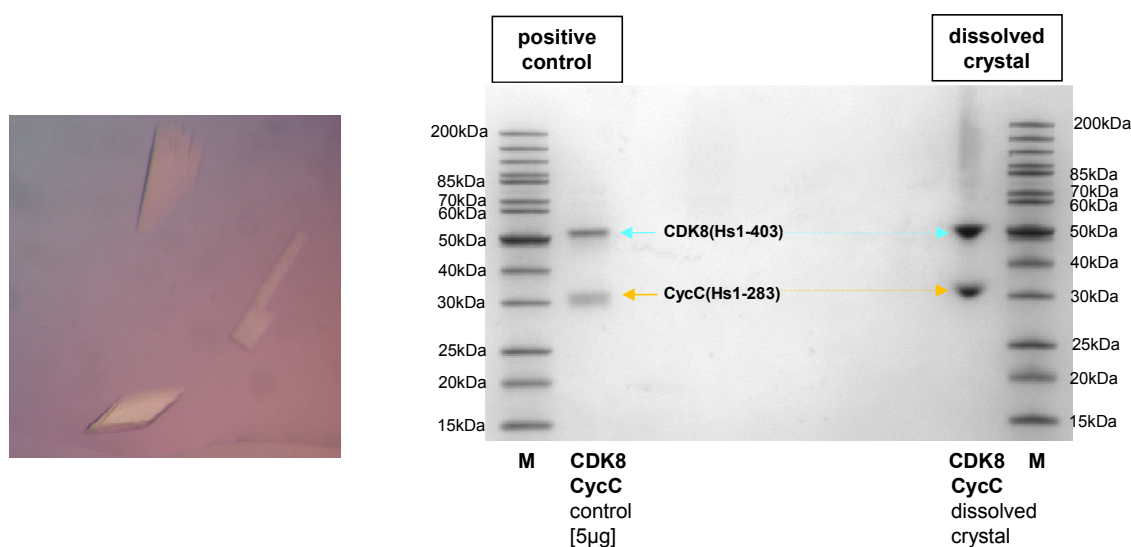


Figure 18 CDK8/CycC/sorafenib protein crystals

left: morphological pattern of CDK8/CycC/sorafenib crystals grown in 20%PEG3350 and 0.2M LiCl; *right*: M= marker; analysis of the crystal by 10% Bis-Tris SDS-PAGE whether the protein complex CDK8/CycC is still formed or influenced disadvantageously by the crystallization conditions

For refinement of the protein crystals, the influence of different additives such as 4% DMSO was assessed. Furthermore the right choice of cryoprotectant was required to vitrify the protein drop under cryogenic nitrogen gas during the mounting on the goniometer head to avoid the formation of ice and a negative influence on the protein crystal itself by damage or degradation (Pflugrath, 2004). All crystals were harvested in mother solution containing different concentrations of cryoprotectant before flash-cooling. Protein crystals then were mounted and analyzed for the formation of ice rings in the preliminary diffraction tests. Moreover the experiment was morphologically judged for degradation or damage on crystals caused by non-suitable cryo reagents. The addition of 20-25% ethylene-glycol (EG) to the mother liquor appeared as the best cryoprotectant of CDK8/CycC crystals for the time being. The refinement of the crystals and the cryo-conditions resulted in the best diffraction pattern of

a CDK8/CycC crystal with a resolution of 2.2Å (see Figure 19). Subsequently a data set was taken at the SLS.

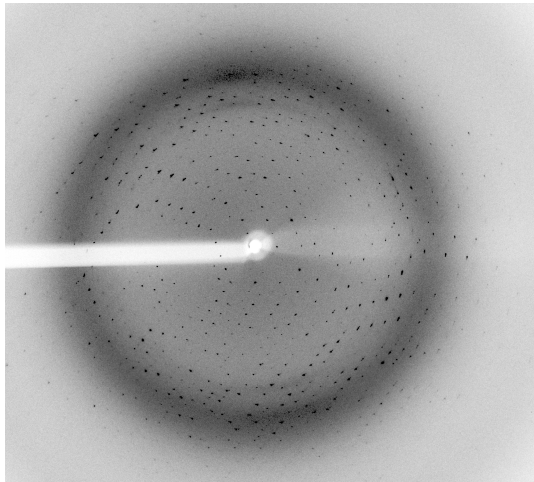


Figure 19 Preliminary test of diffraction with the maximum reasonable crystal to detector distance

3.5 Determination of the crystal structure of the CDK8/CycC/sorafenib complex

Collection of data of the CDK8/CycC/sorafenib complex and its processing was performed as described in material and methods. However, after the molecular replacement using the automatic molecular replacement pipeline BALBES (Long et al. 2008), the R-factors (R_{work} 40.3%; R_{free} 46.4%) displayed surprisingly good values compared to the corresponding electron density. The structure was solved by a combination of manual building and improvement of the electron density by the program Buccaneer (Cowtan, 2006) to avoid a new approach of structure solution using experimental phasing. A Buccaneer-supported approach (Figure 20) was used to improve electron density to provide phase information suitable for manual model building together with a manual correction of the model build by Buccaneer. In more specific terms this means deletion of amino acids which were not defined by electron density as well as insertion of amino-acids which were defined by displayed density. Moreover side-chains were rebuilt and water molecules inserted if possible and stereo-chemically reasonable.

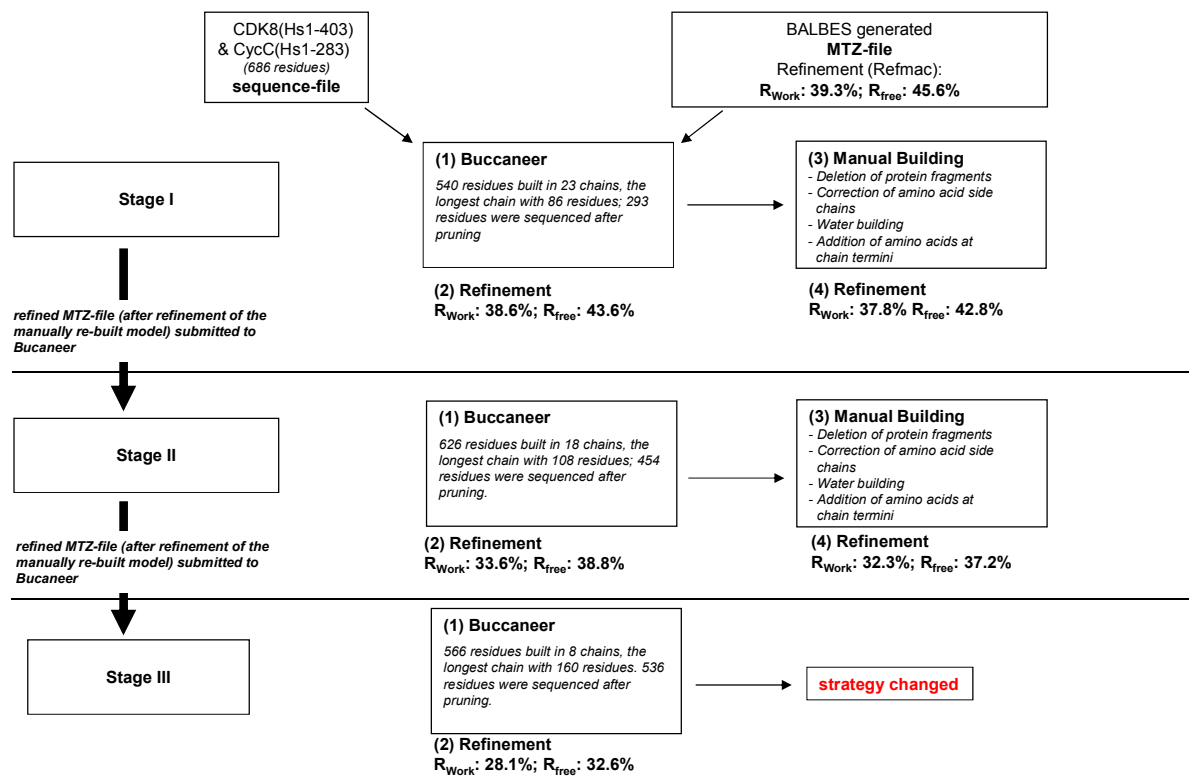


Figure 20 Buccaneer-based approach to improve electron density and increase phase information

After three steps of Buccaneer-based building, the strategy was changed (see Figure 21) as the available phase information has reached a quality that facilitated manual building. Subsequently, the individual eight chains of the last buccaneer modified pdb-file were merged to one. Thereafter, according to alignment with target sequences of CDK8 and Cyclin C, corresponding chains were separated to chain A (CDK8) and B (CycC). Manual model building was now performed without Buccaneer according to the sequence-alignments with subsequent refinement using Refmac. After several refinement-cycles the model comprising CycC (chain B) had reached an optimum. In the CDK8-model (chain A) the CDK8 C-lobe was almost defined by electron density with some segments un-ordered whereas the CDK8 N-lobe was only poorly defined at this stage. Therefore the MTZ-file was again submitted to Buccaneer. After superimposition of both models in COOT, residues corresponding to the N-lobe were merged with the rest of chain A. After 11 cycles of refinement the CDK8/CycC-inhibitor sorafenib was modeled into positive electron density that was unambiguously defined by the $F_{\text{O}}-F_{\text{C}}$ map.

Stage	R-Factors after Refinement (Refmac 5)
01	R _{Work} : 28.6%; R _{free} : 33.2%
02	R _{Work} : 26.6%; R _{free} : 31.2%
03	R _{Work} : 26.2%; R _{free} : 30.4%
04	R _{Work} : 25.3%; R _{free} : 30.0%
05*	R _{Work} : 24.7%; R _{free} : 29.2%
06	R _{Work} : 23.2%; R _{free} : 27.6%
07	R _{Work} : 23.0%; R _{free} : 27.5%
08	R _{Work} : 22.7%; R _{free} : 27.2%
09	R _{Work} : 22.7%; R _{free} : 27.2%
10	R _{Work} : 22.6%; R _{free} : 27.6%
11	R _{Work} : 21.8%; R _{free} : 27.0%
12*	R _{Work} : 21.5%; R _{free} : 26.7%

Annotations***stage_05:**

While CycC (chain B) was complete at this stage, the model of CDK8 (chain A) contained gaps regarding the C-lobe and the N-lobe was only poorly defined. Therefore in a next step the latest modified MTZ-file was again submitted to Buccaneer.

The Buccaneer-built model was superimposed to the latest manually modified model and the residues of the Buccaneer modified model corresponding to the N-lobe were merged with chain A.

***stage_12**

CDK8/CycC-inhibitor Sorafenib had been omitted from the phasing model until now. It was modelled into positive electron density using COOT and is unambiguously defined.

Figure 21 Manual building and refinement cycles leading to an electron density that unambiguously defines the inhibitor within the F_O-F_C map

Final refinement cycles led to a CDK8/CycC model of a high quality with R_{work} 18.1% and R_{free} 22.1% (Table 4).

Data Collection	
Beamline	SLS X06SA
Space group	P 1 21 1
Unit cell dimensions	a=70.7Å b=70.6Å c=79.1Å $\alpha=90^\circ$, $\beta=108.3^\circ$, $\gamma=90^\circ$
Data Processing	
Resolution (Å)	75.13 – 2.2
Average Mosaicity (deg.)	0.81
R _{merge}	0.089 [0.542] ^a
Total number of observations	174390
Total number unique observations	37725
Mean I/sd(I)	10.7 [2.6]
Completeness	100 [99.9]
Multiplicity	4.6 [4.6]
Refinement	
Total number residues ^b	621
Number of water molecules	175
R _{Cryst}	18.1
R _{Free}	22.1
r.m.s. bonds (Å)	0.007
r.m.s. angles (deg.)	1.00
Rachmandran plot	
Most favoured geometry (%)	93
Additionally allowed (%)	7

Table 4 Crystallographic parameters for human CDK8/CycC/sorafenib (3RGF)

^a Values in brackets refer to the highest resolution shell ^b The final model compromises residues for CDK8(-1-353) and CycC(-1-264) as two additional residues were introduced (termed -1, 0 whereas 1 refers to Met1^{CDK8}/Met1^{CycC} according to sequence in Swiss-Prot-ID) by the N-terminal expression tag. Segments, which are not defined in electron density, include CDK8 residues 27-32, 115-123, 179-195, 238-245 and 354-403 and CycC residues 265-283.

4. The first structure of CDK8/CycC within the CDK8-DMG-out conformation implicates specificity within the CDK/Cyclin family and reveals interaction with a deep pocket binder

4.1 The Overall Structure of CDK8/CycC

The overall CDK8/CycC architecture (Figure 22, above) is related to both, the previously solved CDK9/CycT and CDK2/CycA structures, where the binding mode of the transcriptional CDK9/CycT pair differs from the cell cycle complex CDK2/CycA by a relative rotation of the two components by 26° (Baumli et al., 2008). An intermediate arrangement between these extreme positions is found for CycC when superimposing CDKs. CycC is rotated by about 10° along its H4 helix (Figure 22, below left) relative to CycA while ~15° are still missing towards the CycT position in the CDK9/CycT complex (Figure 22, below right).

Similar to the other CTD kinase complex CDK9/CycT, CycC almost exclusively contacts the CDK N-lobe while CycA contacts both the N-lobe and the C-lobe in the CDK2 complex. Unexpectedly, the overall contact surface between CDK8 and CycC is 1614Å² but only 970Å² in CDK9/CycT (3BLH) (all surface calculations were performed using CCP4 Structure analysis, Dodson et al., 1997). Thus, the N-lobe CDK8/CycC interaction has approximately the same contact area as the CDK2/CycA complex (1689Å²) where both CDK lobes contribute to cyclin binding.

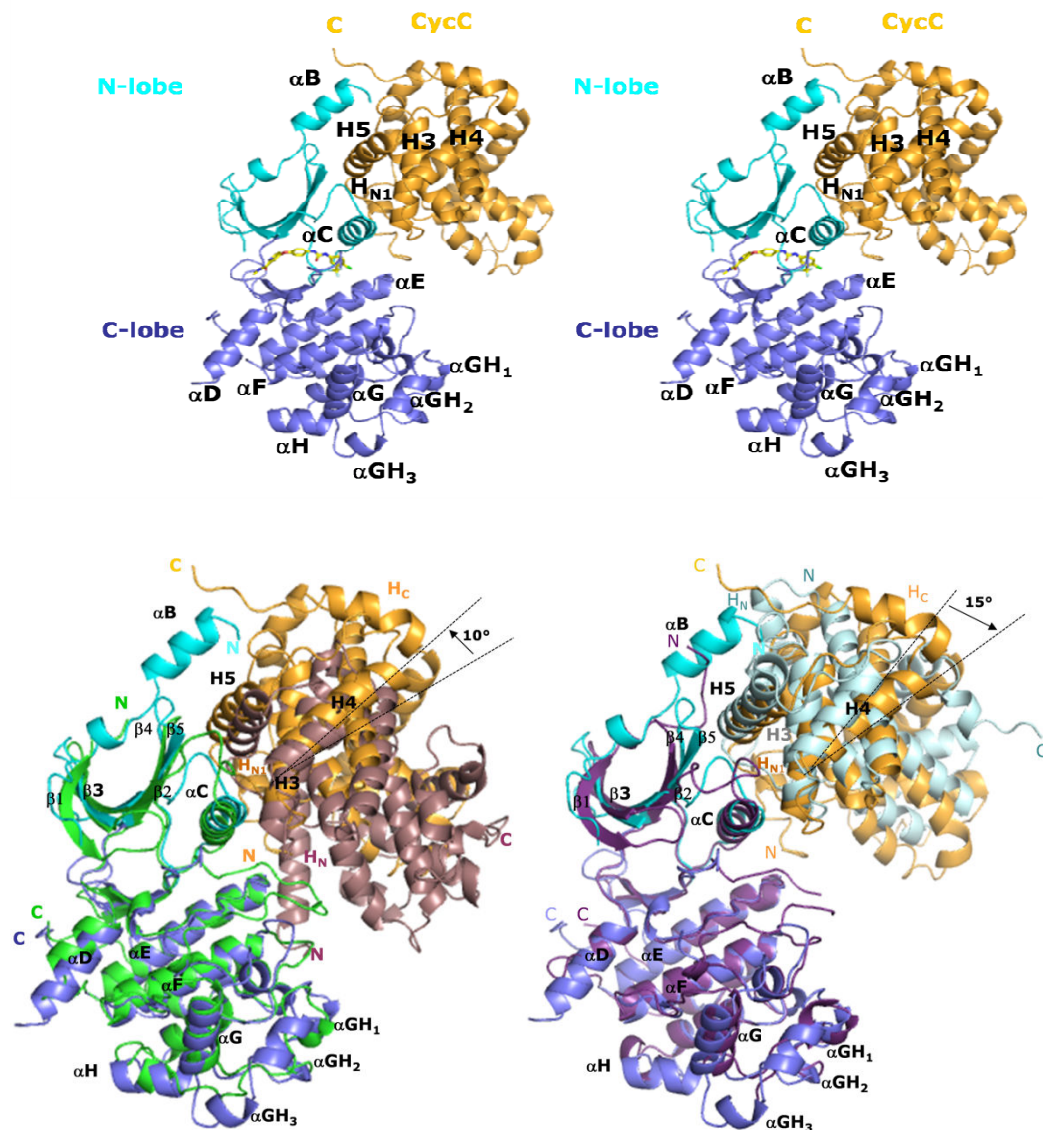


Figure 22 The Overall Crystal Structure of the human CDK8/CycC Protein Complex.

Above: human CDK8/CycC complexed with Sorafenib; CycC (*gold*), CDK8 N-Lobe (*turquoise*), CDK8 C-Lobe (*slate blue*); **Below, left:** Comparison of CDK8/CycC binding manner to CDK/CycA (1QMZ). Both the CDK2 N-lobe as well as the C-lobe are contacting CycA. Superimposition is centred on CDK8 core rmsd 1.60Å, sequence identity 38.5% (COOT); CDK2 *green*, CycA *dusky pink*; **Below, right:** Comparison of CDK8/CycC binding manner to CDK9/CycT (3BLQ). The binding surface of CycT is limited to the N-lobe similar as in CDK8/CycC. However, the H4 helices are misaligned in CycC/CycT by 15°. The overlay is centred on CDK8 core rmsd 1.56 Å; sequence identity 35.8% (COOT); CDK9 *yellow*, CycT *light blue*

4.2 Crystal Structure of human CycC

The human CycC crystal structure adopts the general cyclin box fold (CBF) (Noble et al., 1997). The highly conserved CBP includes two canonical repeats each containing 5 helices encircled by an N-terminal helix (H_N) and a C-terminal helix (H_C), respectively. Those are highly conserved among the cyclin family as underlined in the structural alignment of CycC with SRB11, CycT1, and CycA (Figure 23). However, there are differences in the orientation, length and number of the H_N and H_C helices (Figure 23). The CycC *Saccharomyces pombe* analogue SRB11 (1ZP2) and human CycC can be superimposed closely regarding the central elements (core rmsd 1.65Å and 31.4% sequence identity using COOT, Figure 24). Striking differences are observed in the N- and C-terminal region of CycC/SRB11. The CycC N-terminus including the comparatively short CycC H_{N1} helix contacts its binding partner CDK8 whereas the longer CycC H_{N2} helix is beyond the binding surface with CDK8. This is in contrast to the N-terminus of SRB11, which is pointing in opposite direction (Figure 24) (Hoepfner et al., 2005). However, the SRB11 H_N helix has been postulated to be mobile and its orientation might thus result from the crystal packing forces or crystallization conditions (Hoepfner et al., 2005). In the C-terminal region CycC displays a C-terminal helix and establishes further contacts with CDK8. Since SRB11 is about 35 amino acids shorter than CycC, it lacks a C-terminal helix and ends after the H_5' helix. Hoepfner et al. described a highly conserved groove between the two cyclin repeats in their crystal structure of the CycC analogue SRB11. This groove is exclusively conserved within CycC family members (Hoepfner et al., 2005). Five surface residues of this groove are invariantly conserved (I42^{CycC}/ I32^{SRB11}, R58^{CycC}/ R49^{SRB11}, W177^{CycC}/ W177^{SRB11}, D182^{CycC}/ D165^{SRB11}, Y184^{CycC}/ Y164^{SRB11}; Figure 24) with their nature deviating strongly from other cyclins (Figure 23). However, they are far too distant (~22Å) from CDK8 to be relevant for CDK interaction (Figure 24). They may be involved in either binding of MED12 and MED13 or substrate recognition. The CycC specific groove could play an analogous role as the conserved surface patch in CycA that binds kinase substrates (Schulmann et al., 1998) and may bind, for example, the peptide repeat Y₁S₂P₃T₄S₅P₆S₇ of RNA Pol II CTD (Meinhart et al., 2005). Corresponding surfaces on CycA, CycE, CycD and CycB differ in their electrostatic characteristics and may thereby also account for substrate specificity (Lolli, 2010).

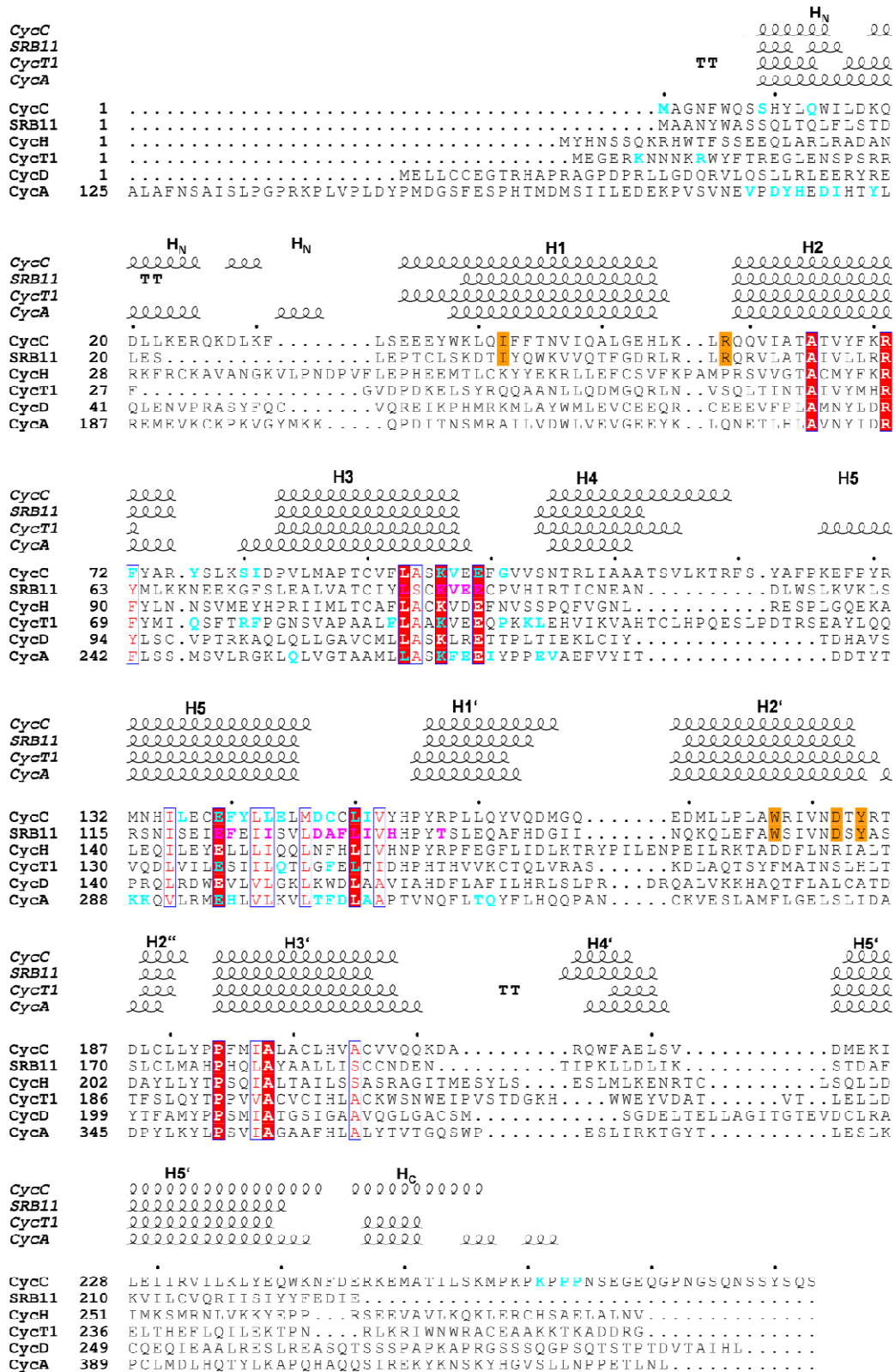


Figure 23 Cyclin C (legend on next page)

Figure 23 Cyclin C

Comparison of human CycC sequence with its yeast analogue SRB11 (1ZP2) and other human cyclins. Sequence alignment prepared with ClustalW (Larkin et al., 2007), the secondary structure alignment with ESPript (Gouet et al., 1999). The nomenclature of secondary structure elements is adapted from Hoepfner and co-workers (Hoepfner et al., 2005). Identical residues are red boxed, similar ones (>90%) are coloured red and framed blue. Residues conserved in CycC species are boxed orange. Human CycC residues interacting with human CDK8 (distance $\leq 4\text{\AA}$) are coloured turquoise. For comparison SRB11 residues are coloured pink that were supposed to interact with yeast CDK8 (Hoepfner et al., 2005). Residues in CycT1 and CycA below a distance of $\leq 3.5\text{\AA}$ are coloured turquoise according to Baumli and co-workers (Baumli et al., 2008). The 3RGF CycC structure ends with Pro264^{CycC}.

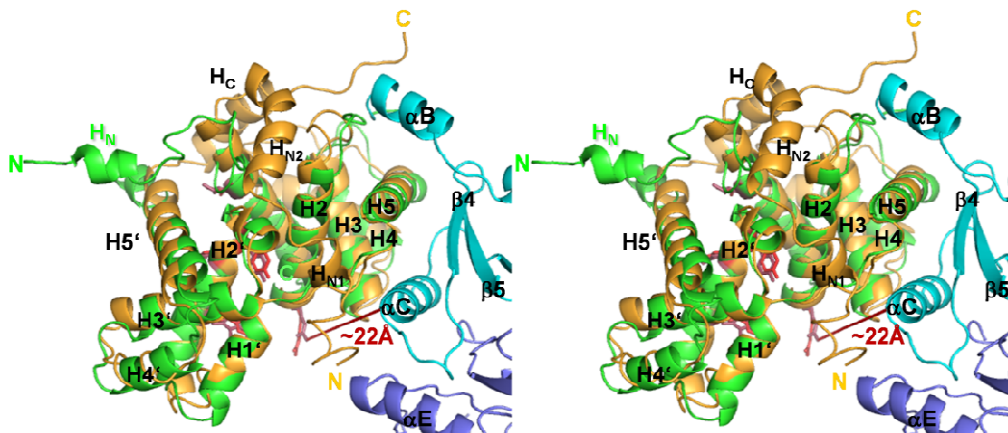


Figure 24 Cyclin C

Superposition of CycC with SRB11. CycC (gold), Srb11 (green), N-lobe CDK8 (*turquoise*), C-lobe CDK8 (*slate blue*); for clarity protein complex was rotated approximately 180° relative to Figure 23. Only parts of CDK8 that are near to interaction surface with CycC are shown for clarity as well. Five residues, highly conserved only amongst CycC species, are marked in red (CycC: I42^{CycC}, R58^{CycC}, W177^{CycC}, D182^{CycC}, W184^{CycC}) or salmon red (SRB11: I33^{SRB11}, R49^{SRB11}, W160^{SRB11}, D165^{SRB11}, W167^{SRB11}) and too distant to form contacts to CDK8 ($\sim 22\text{\AA}$).

4.3 Crystal structure of human CDK8

CDK8 shows the typical bilobal kinase fold consisting of the N-lobe (residues 1-96) and the C-lobe (residues 97-353) with the deep catalytic cleft between the two lobes (Bossemeyer et al., 1993, Huse et al., 2002). In the secondary structure alignment (Figure 25), CDK8 exhibits the core structure elements that are very well conserved for all family members in the common range of differences in lengths (Lolli, 2010). Structural deviations include the reduction of the β 1-strand (Figure 22 and Figure 25) to a minimal length of two amino acids. The α D helix is elongated only in CDK8 by a positive amino acid cluster, which extends in the loop region between α D and α E (Figure 22 and Figure 25). Several residues in α D- α E region are not included in electron density and may need to be stabilized by binding of an interaction partner. In this respect the α D helix and the α D- α E loop were hypothesized to serve as mutual recruitment sites in a CDK7/CDK2 binding model (Lolli et al., 2007). Binding of MED12 or MED13 to the α D region could influence the kinase activation process due to the proximity to the active site. Other structural peculiarities cluster close by such as the CDK8 specific helices α GH₁₋₃. Additionally, there is a CDK8-unique insertion consisting of nine residues (²⁴⁰EDIKTSNPY²⁴⁸) in front of the α G helix. Helix forming residues in the α G- α H region are highly conserved among CDK8-species (CDK8 species alignment). In accordance to the here observed structural elements of CDK8 the α G helix and the α G- α H loop in general constitute a site for protein-protein interactions in the CDK-family (Lolli, 2010). In the N-lobe there is an additional N-terminal helix α B preceding the prominent α C helix (Figure 22). This was already obvious in the secondary structure prediction and plays a pivotal role in the cyclin recognition as will be discussed in the following paragraphs. Other CDKs are shorter by 5-10 residues in the CDK8 α B helix region (Figure 25) except for CDK9, which shows an equal length, but the CycT1 complex structure shows only a short helical turn at this site (Figure 22 and Figure 25).

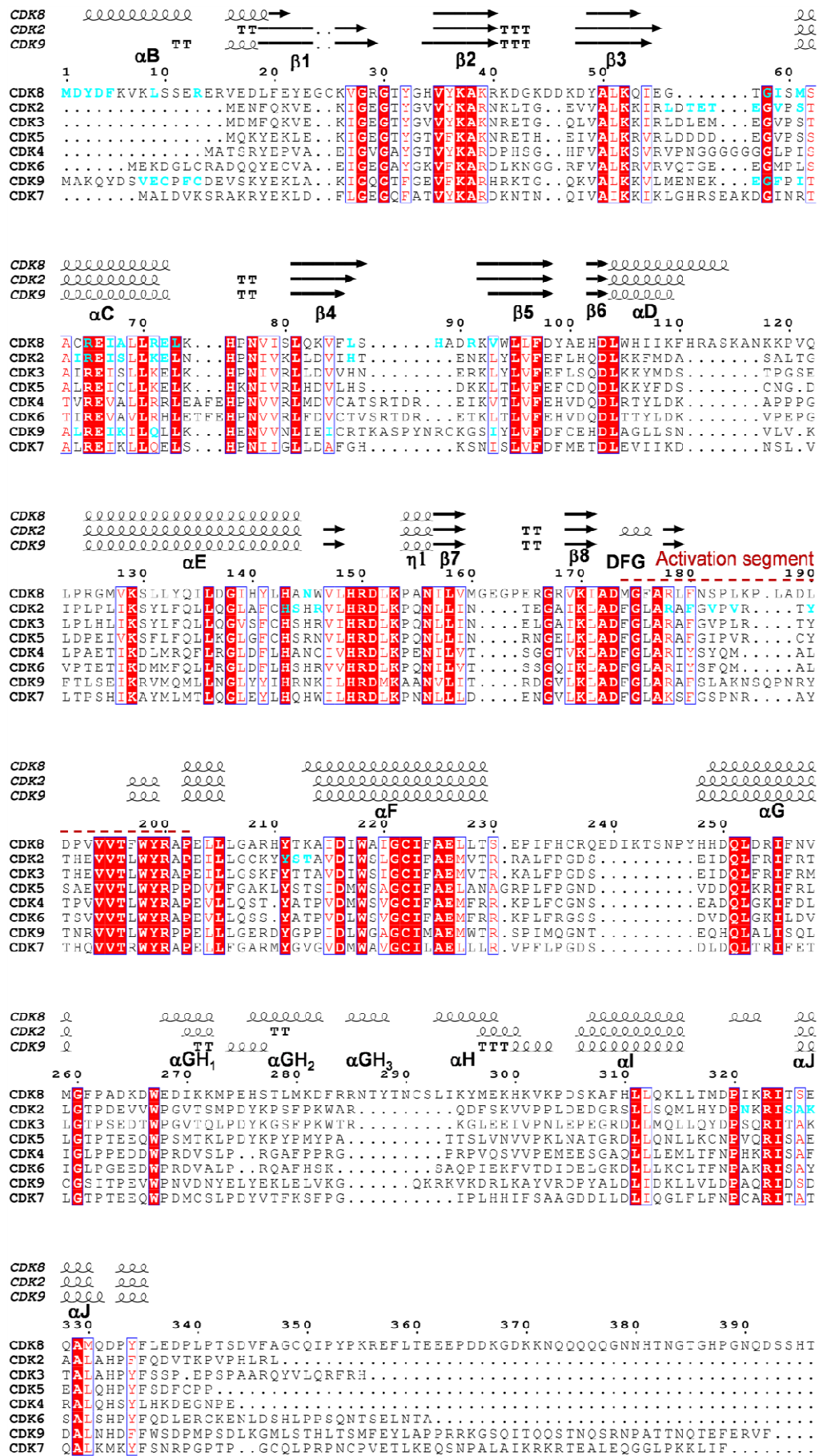


Figure 25 Secondary Structure alignment of CDK8 within CDK family members (legend on next page)

Figure 25 Secondary Structure alignment of CDK8 within CDK family members

Sequence alignment was prepared using ClustalW (Larkin et al., 2007), alignment of secondary structure using ESPript (Gouet et al., 1999). Nomenclature of secondary structure elements is according to Lolli, 2010. The secondary structure for CDK8, CDK2 and CDK9 is indicated above. Identical residues are red boxed, residues with an identity $90\% <$ are coloured *red* and framed *blue*. Most strikingly CDK8 displays an additional N-terminal helix α B. The α D helix is elongated by a cluster of positive amino acids, which extends in the α D- α E loop. There are additional secondary structural elements within the α G and α H region, termed α GH₁, α GH₂ and α GH₃. Sequential deviations include an insertion of nine residues in front of the α G-helix. CycC interacting residues in CDK8 within a range of 4Å are colored *turquoise*. Cyclin interacting residues in CDK2 and CDK9 are colored *turquoise* as well according to Lolli, 2010. CDK9 maintains N-terminal interactions with CycT similar to the CDK8/CycC complex. The common interaction surface on the prominent α C helix is highly conserved. Additional contacts in CDK8 to the cyclin cluster in the region between the β 4- β 5 loop. In contrast to CDK9/CycT the residues within the CDK2 C-lobe contact CycA as well according to its binding manner. CDK8 maintains a sparse interaction to CycC within the α E helix. The 3RGF CDK8 structure ends with Tyr353^{CDK8}.

4.4 The conserved core of the CDK/cyclin interface of CDK8/CycC: comparison with CDK9/CycT1

The center of the interaction of the novel CDK8/CycC complex is formed by a conserved core (Figure 26) as in other CDK/cyclin pairs (Lolli, 2010; Echaliier et al., 2010).

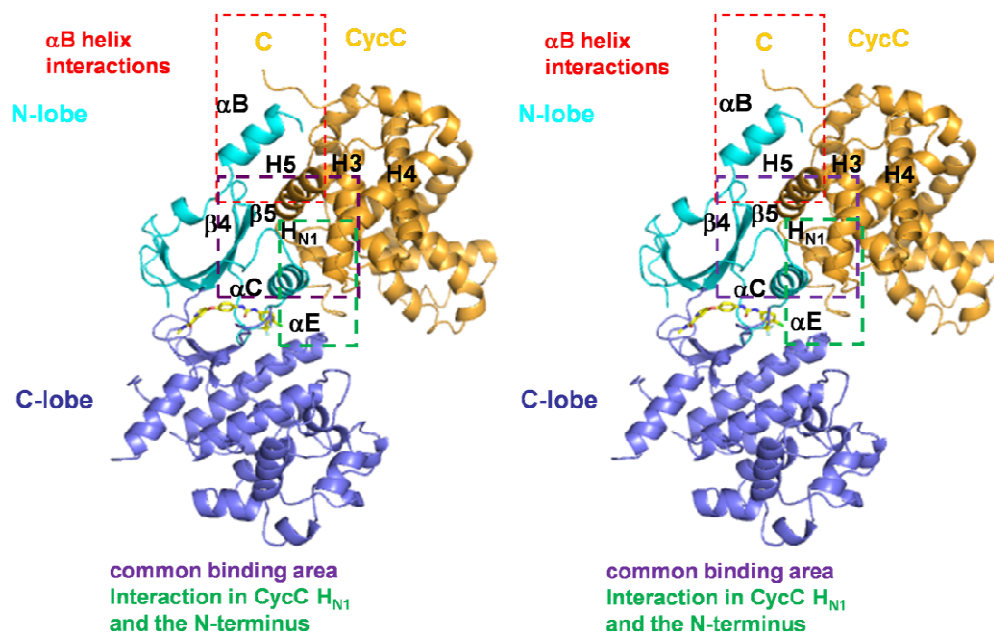


Figure 26 CDK8/CycC binding surface.

CDK8/CycC overall structure with CycC binding sites; CycC (*gold*), CDK8 N-lobe (*turquoise*), C-lobe (*slate blue*); Besides the common binding surface the CDK8/CycC complex maintains additional contacts mediated by the CDK8 N-terminal α B helix and the contacts with the CycC N-terminus and H_{N1} .

In detail, the CycC H5 helix contains the central hydrophobic residue (F140^{CycC}), which functions as a kind of “pivot point” for the interaction surface (Figure 27, above). In addition, the C-terminal end of CycC H5 establishes an intimate contact with CDK8 (Figure 27, above). The orientation of the CycC H5 helix seems to be defined by interaction with both the N-terminal CDK8 α B helix (as will be discussed later) and the common binding surface. In CDK8/CycC (Figure 27, above) F140^{CycC} (H5) is bound by the CDK8 β 4- β 5 region (L86^{CDK8}, H88^{CDK8}, R91^{CDK8} and V93^{CDK8}) and further embedded in CycC H5 helix residues.

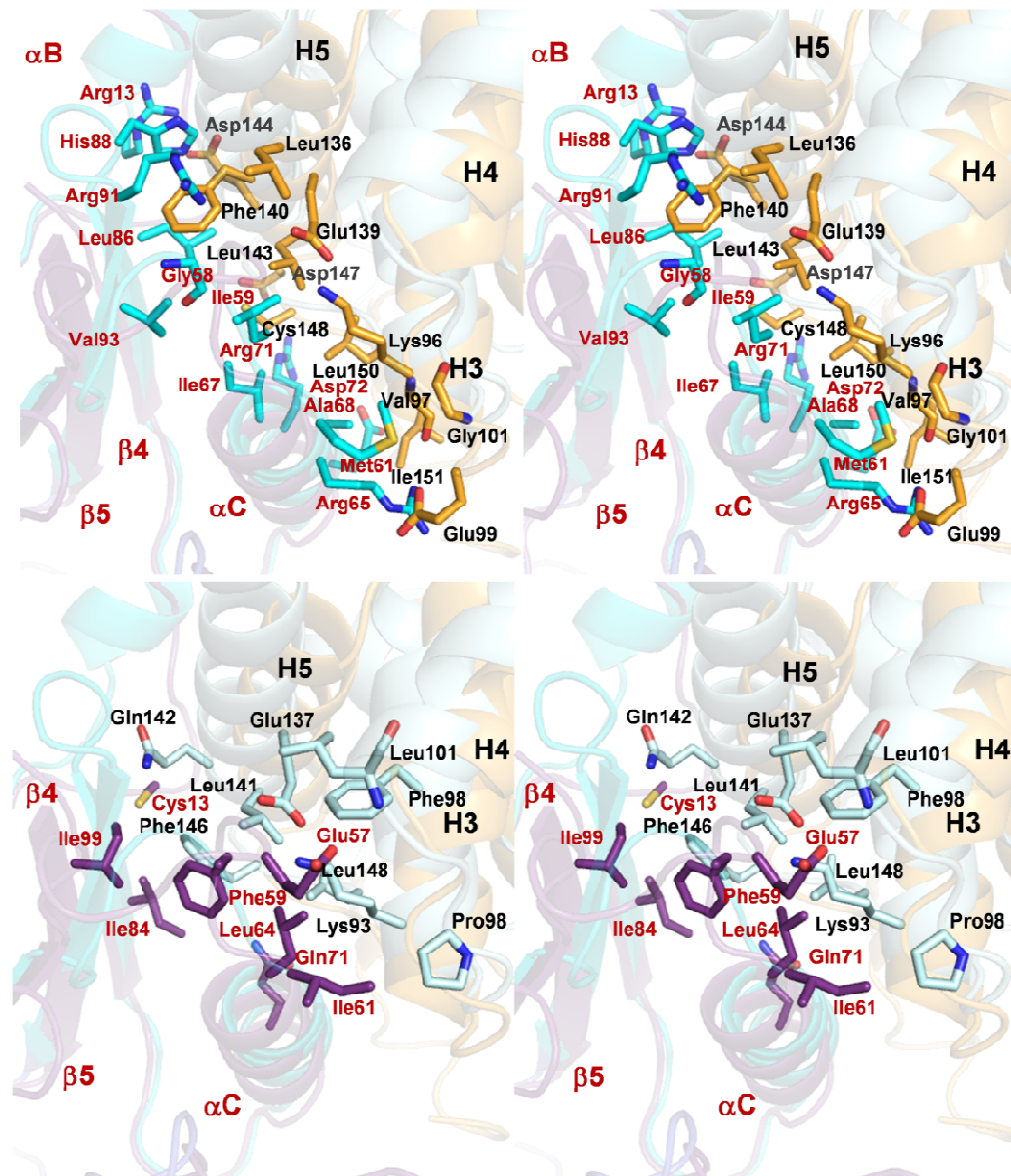


Figure 27 CDK8/CycC binding surface in comparison to CDK9/CycT1 binding surface.

CDK8/CycC overall structure with CycC binding sites; CycC (*gold*), CDK8 N-lobe (*turquoise*), C-lobe (*slate blue*); Besides the common binding surface the CDK8/CycC complex maintains additional contacts mediated by the CDK8 N-terminal α B helix and the contacts with the CycC N-terminus and H_{N1} . Close Up of the common binding surface in CDK8/CycC and CDK9/CycT. CDK9/CycT (3BLQ) superposed to CDK8/CycC. CDK9 (3BLQ; violet) was centered on CDK8 (N-lobe turquoise, C-lobe pale blue) with core rmsd 1.56Å and a sequence identity of 35.85% (COOT) to make differences in the cyclin binding mode more obvious (CycC (*gold*), CycT (*metallic*)). Interacting residues are coloured according to atom type (oxygen *red*, nitrogen *blue*, sulfur *yellow*). CDK residues and structural element are labelled *red*, those of the Cyclin in *black*.

The CDK9/CycT complex (Figure 27, below), like most others, differs in that the central cyclin residue F146^{CycT} is not part of H5 but of the following turn (Echalier et al., 2010). In CDK9/CycT, like

in most other complexes, the pocket is formed by residues of the α C helix. In CDK8/CycC, however, the β 3- α C region (G58^{CDK8}, I59^{CDK8}) is shifted towards the CycC H5 helix (E139^{CycC}) thereby enabling additional contacts not present in CDK9. Besides the interaction in the β 3- α C region, several contacts along the CDK8 α C helix (R71^{CDK8}, D72^{CDK8}, A68^{CDK8}, R65^{CDK8} and M61^{CDK8}) with the CycC H3 helix (G101^{CycC}, E99^{CycC}, V97^{CycC} and K96^{CycC}) and the H5-H1' region (I151^{CycC}, L150^{CycC}, C148^{CycC} and D147^{CycC}) maintain the CDK8 α C helix in its active “in-conformation” (Jeffrey et al., 1995; Pavletich, 1999). These CDK8/CycC interactions correspond to the common interaction surface including the β 3- α C region, α C-helix and post- α C region (β 4- β 5) on CDKs (Lolli, 2010; Echali er et al., 2010). Besides to the CDK8/CycC specific variations of the common binding surface (Figure 27), specificity is achieved by additional binding sites (Figure 26) as described in the following paragraphs.

4.5 The adaptable H_N helix of CycC is contacting CDK8

The aforementioned CDK8/CycC contacts that are suggested to rearrange the CDK8 α C helix and move it towards the activation segment seem to be stabilized by interaction with the CycC N-terminus (Figure 28).

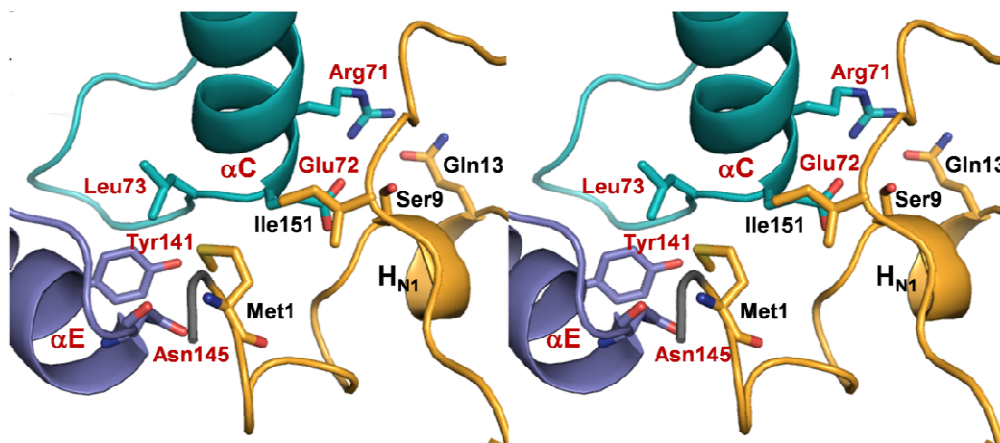


Figure 28 The N-terminus of CycC forms contacts with CDK8.

CycC (gold) interacting residues with CDK8 N-Lobe (turquoise) respectively C-lobe (slate blue); Atoms are coloured according to atom type (nitrogen dark blue, oxygen red, sulfur yellow). Artificial residues (main-chain) from the expression-tag are coloured in dark grey. CDK8 residues and structural element are labelled red, those of CycC black.

The side-chain of M1^{CycC} is observed in van der Waals contact distance with Asn145^{CDK8} and Y141^{CDK8} (α E helix) and L73^{CDK8} located in the α C helix. The interaction is completed by a hydrogen bond between the main-chain nitrogen of M1^{CycC} and the main-chain carbonyl of N145^{CDK8}. Hydrogen bonding between CDK8 α C helix and CycC H_{N1} helix via E72^{CDK8} and the main-chain carbonyl of S9^{CycC} and the side-chains of R71^{CDK8} with Q13^{CycC} further stabilizes the interaction. In short the CDK8 interactions involve the N-terminus of CycC and the first H_{N1}-helix. These elements stabilize the α C helix and involve the CDK8 C-lobe. This observation is contrary to published literature where the H_N helix of a transcriptional cyclin was postulated to be located beyond the binding surface and only the C-terminal H_C helix expected to contact the CDK (Lolli, 2010). This is true for CycK, CycH and CycT even though the N-terminus of CycT maintains several contacts with CDK9.

4.6 CDK8 contains an additional N-terminal CycC recognition helix

The N-terminal α B helix of CDK8 is unique within CDKs and maintains contacts with the CycC H5 helix in an unprecedented way (Figure 29).

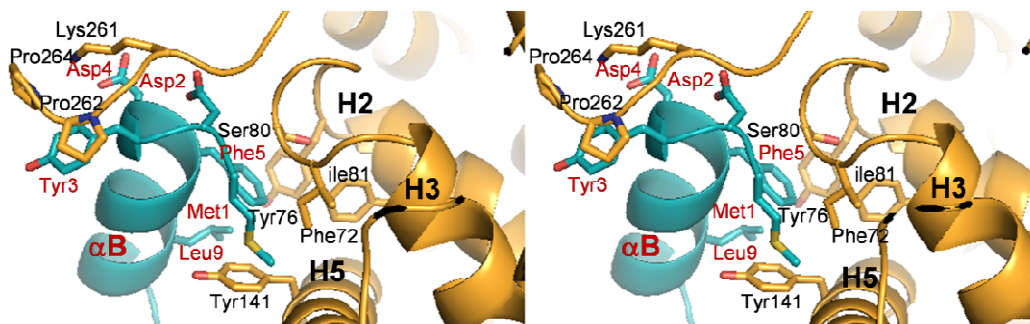


Figure 29 CDK8 α B helix interacting with CycC.

CycC (gold) interacting residues with CDK8 N-Lobe (turquoise); Atoms are coloured according to atom type (carbon yellow, nitrogen dark blue, fluorine grey, chloride green). Artificial added residues from tag are colored dark grey. CDK8 residues and structural elements are labelled in red, those of CycC in black. The electron density for D-1^{CDK8} and K0^{CDK8} is very weak. Only main chain is observed in electron density and is considered unlikely to have any influence on binding surface with CycC.

F5^{CDK8} packs against Y141^{CycC} and is additionally shielded from the solvent by L9^{CDK8}, F72^{CycC}, Y76^{CycC}, S80^{CycC} and I81^{CycC} and more distantly by M1^{CDK8}. Further electrostatic contacts between the side-chains of D2^{CDK8} and D4^{CDK8} stabilize the orientation of the main chain amine of K261^{CycC} by

main-chain and side-chain interactions. Contacts between R13^{CDK8} and E144^{CycC} (not shown) complete the interaction. The CDK8 α B helix also forms a hydrophobic interaction based on Y3^{CDK8} involving the CycC Pro-rich C-terminal region (P263^{CycC}, P264^{CycC}). All charged residues of the N-terminus of CDK8 as well as the CDK8-contacting residues in the CycC C-terminus are species-specific (see structure alignments Figure 23 and Figure 25).

Based on the structure, an N-terminally shortened CDK8-construct (Hs23-464) was designed in an attempt to understand the contribution of the α B-helix in the specific binding to CycC. In the absence of sorafenib, K_d values of mutant and native CDK8/CycC complexes were determined (Figure 30, chapter 2.14).

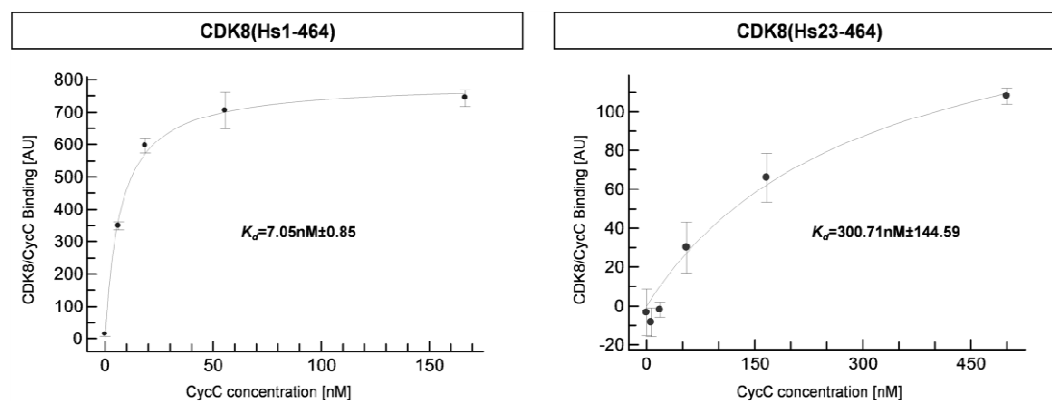


Figure 30 K_d -Determination for the interaction between full-length CDK8(Hs1-464) with CycC(Hs1-283) and between the α B-helix shortened CDK8(Hs23-464) with CycC(Hs1-283) in absence of sorafenib.

CycC affinity for both CDK8 constructs [2.2nM]; full-length CDK8(Hs1-464) (left) or N-terminal shortened CDK8(Hs23-464) (right); K_d is calculated by non-linear regression as fitted accordingly to equation (1). Affinity for CycC is significantly decreased in the N-terminal shortened CDK8 variant.

The observed K_d values for CDK8(Hs1-464) with CycC are 7.4nM and reduced to 297nM for CDK8(Hs23-464), respectively. Some native CDK/Cyclin affinities have been reported before which are weaker by at least one order of magnitude. CDK9/CycT1 showed a K_d of approximately 300nM (unpublished data by Baumli et al., 2008) whereas 52nM and 57nM were determined for CDK2/CycA and CDK7/CycH respectively (Heitz et al., 1997). Importantly, the stability of the physiologically relevant complexes CDK7/CycH and CDK2/CycA is 25-50 fold higher than the 'unnatural' complexes

CDK2/CycH or CDK7/CycA respectively (Heitz et al., 1997). Thus, deletion of the α B helix in CDK8 reduces the specificity towards CycC by a similar factor. The residual binding affinity of the CDK8/CycC pair in absence of the α B helix, similar to CDK9/CycT1, is not surprising as the common binding surface is still present. The K_d measurements are consistent with a large interaction surface area. In the absence of the α B-helix, the interaction surface area is dramatically reduced from 1614Å² to 948Å², a value seen in CDK9/CycT (Echalier et al., 2010). CycC promiscuously interacts with CDK8 and CDK3 (Casamassimi et al., 2007; Ren et al., 2003), which has no α B helix and binds very likely less tightly.

In summary, the here presented structural and mutational data explains the exceptionally high affinity between CDK8 and CycC. In addition to discriminating of binding against CDK3, the high affinity of CDK8/CycC could also explain selectivity amongst transcriptional CDKs and their cognate cyclins. The overall binding of CDK9/CycT is quite similar to CDK8/CycC but significantly less tight. Very tight binding to a specific partner is one mechanism to gain specificity within candidate interaction partners similar in sequence or structure (Schreiber et al., 2011) and may be achieved by small structural variations as seen here in the CDK8 CycC recognition helix α B, which discriminates amongst transcriptional CDK/Cylin pairs. Moreover, as part of a multi-protein assembly, a very stable CDK8/CycC complex may be required for the specificity of the recruitment of the other CDK module components MED12 and MED13, that were shown to associate independently of each other (Knuesel et al., 2009b).

A recent publication highlighted the impact of CycC competing inhibitors with CDK8 or CDK3 (Rajender et al., 2011). The authors virtually screened inhibitors against a human CycC model that may disable contacts within the common binding surface. Based on the here presented mutational experiments, the specific CDK8 N-terminus cyclin interaction appears as target for interfering with the CDK8 function as transcriptional regulator.

4.7 Sorafenib binds to the hinge region in CDK8 inducing DMG-out conformation

The inhibitor sorafenib binds to the CDK8 hinge region within the catalytic cleft between the two kinase lobes with a buried surface of 92.1%. It is unambiguously defined in electron density (Figure 31), but was included only in the final steps of refinement.

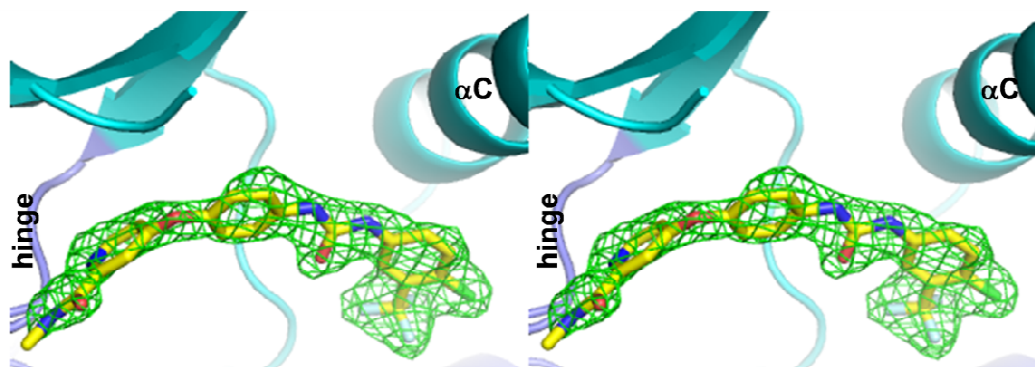


Figure 31 Sorafenib Binding Mode.

F_0 - F_c omit map for sorafenib complexed to CDK8/CycC with contour level of 3σ (green). Atoms are coloured according to atom type (carbon on yellow, nitrogen in blue, chloride in green). The inhibitor had been omitted from the phasing model, which had been refined several cycles before modelling in electron density. Sorafenib binds to the CDK8 hinge region in the catalytic cleft between the two lobes.

Sorafenib induces the so-called DFG-out conformation in kinases, which is actually DMG-out in CDK8 as exception to all other family members (Figure 25). The DMG-conformation of CDK8/CycC is similar to the DFG-motif observed in the crystal structure of CDK6 complexed with the tumor suppressor inhibitor of CDK4 and CDK6 (INK4) (1BI7, 1BI8, Russo et al., 1998; 1GN3, Jeffrey et al., 2000). Although INK4 has a completely different binding mode, opposite to the cyclin binding site, and no small inhibitor molecule present, the DMG-/DFG-triads of CDK8/CDK6 can be aligned closely. However, the catalytic arginines R178^{CDK8} and R168^{CDK6} in the T-loop are displaced by 13.1 Å. This is explainable by comparing the mode of how DMG/DFG-out conformation in both CDKs is achieved. In CDK6, the binding of INK4 adjacent to the catalytic cleft induces large conformational changes. Thereby the CDK6 β 1- β 2 loop is translocated towards the activation segment and the CDK6 N- and C-lobes get misaligned by a 15° rotation (Russo et al., 1998), resulting in an α C helix that is either disordered (1BI7, 1BI8) or misplaced (1GN3). On the other hand the orientation of the T-loop in

CDK8 seems to be determined by contacts with sorafenib. In the close-up of sorafenib bound to CDK8 (Figure 32) the urea linker within sorafenib forms two hydrogen-bonds with E66^{CDK8}, which is a residue absolutely conserved in the α C helix of kinases. As a result, the kinase-wide conserved salt-bridge between E66^{CDK8} and K52^{CDK8} involved in ATP-binding within the active kinase state (Bossemeyer et al., 1993; Huse et al., 2002; Jeffrey et al., 1995) cannot be established. Instead K52^{CDK8} points towards the main chain carbonyl of M174^{CDK8} within hydrogen bonding distance of 3.1Å. Moreover, sorafenib forms one hydrogen-bond to D173^{CDK8} (backbone amide) in the DMG-triad. Sorafenib additionally interacts with the CDK8 hinge region via a hydrogen bond formed by the sorafenib pyridine nitrogen with the A100^{CDK8} main-chain nitrogen and the nitrogen in the N-methyl-4-phenoxy-picolinamid moiety with the A100^{CDK8} backbone carbonyl. The phenyl-moiety of sorafenib is stacked between the side-chains of F97^{CDK8} and M174^{CDK8} and further interacts laterally with the hydrocarbon part of K52^{CDK8} stabilizing its orientation relative to the DMG motif. The 3-trifluoromethyl-4-chlorophenyl ring points into the deep pocket and establishes hydrophobic interactions with F176^{CDK8}. The leucine present at the corresponding position present in all other CDKs is solvent exposed in the observed DFG-in conformations of CDKs.

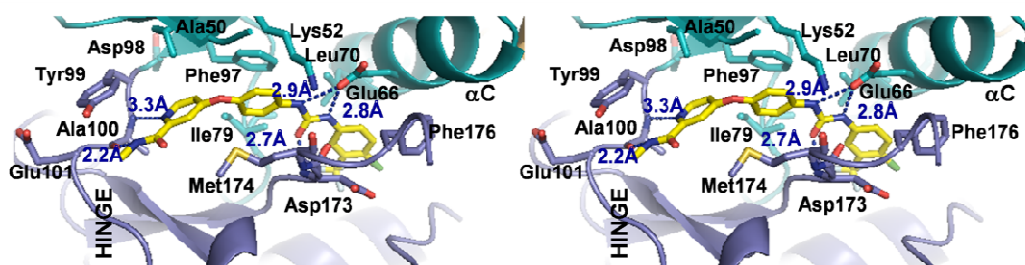


Figure 32 Close-Up of the sorafenib binding mode

Close-Up of the binding mode of Sorafenib to the CDK8 hinge region including H-bond formation (carbon *yellow*, oxygen *red*, nitrogen *dark blue*, chloride *green*, fluorine *grey*). Residues in a distance $\leq 4\text{\AA}$ to Sorafenib are shown as well and coloured according to atom type (oxygen *red*, nitrogen *dark blue*, sulfur *gold*).

Thereby the affinity of CDK8-CycC/sorafenib ($IC_{50}=0.13\mu\text{M}$) was reproduced similar to previously published data ($IC_{50}=0.030\mu\text{M}$) (Neumann et al., 2011). This value is also close to the value obtained by Namboodiri and co-workers for the p38 α kinase ($IC_{50}=0.057\mu\text{M}$, Namboodiri et al., 2010). In the following the binding modes of both kinases will be compared in structural terms (0, 3HEG).

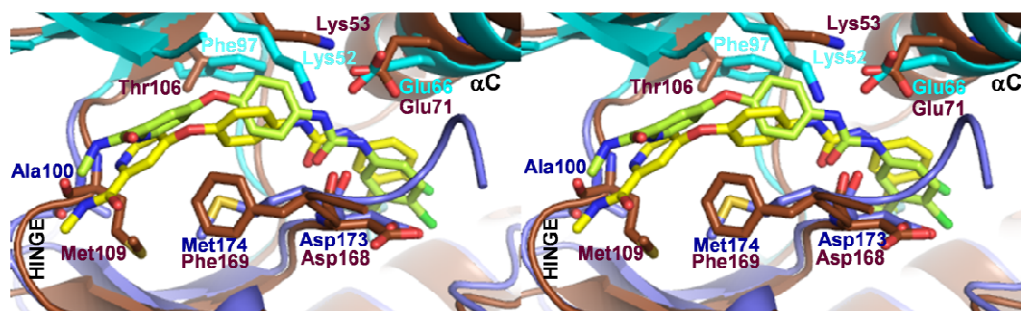


Figure 33 Close-Up of the Sorafenib binding mode of p38 α and CDK8 p38 α

Close-Up of the Sorafenib binding mode of p38 α and CDK8 p38 α (3HEG, brown) complexed with Sorafenib (carbon light green, nitrogen dark blue, oxygen red, chloride green, fluorine grey) superimposed with CDK8 (C-lobe slate blue, N-Lobe turquoise) complexed with Sorafenib (carbon yellow, oxygen red, nitrogen dark blue, chloride green, fluorine grey). Core rmsd 1.92Å, sequence identity 35.4% (COOT). H-bond performing residues and the residues from the DMG/DFG motifs (oxygen red, nitrogen dark blue, sulfur gold) are shown as well as K52^{CDK8}/K53^{p38 α} and F97^{CDK8}/T106^{p38 α} .

Overall, the binding mode of sorafenib is similar. However, flexibility of the hinge region of p38 α allows a different arrangement of sorafenib as observed in a second structure of p38 α /Sorafenib (3GCS, not shown) caused by a rotation of the M109^{p38 α} side-chain out of the nucleotide binding site (Simard et al., 2009a). This binding mode of Sorafenib to CDK8/CycC also is in accordance with general features of DFG-out binders as analyzed by Dietrich et al. in a comparison of sorafenib, imatinib and BIRB796 (Dietrich et al., 2010). The here presented experiments provide structural evidence that CDK8/CycC is a further target of sorafenib, which extends into the deep pocket of the kinase. The recent clinical success of sorafenib and another DFG-out binder, Gleevec[®] (imatinib), was attributed to their deep pocket binding mode (Dietrich et al., 2010). The described sorafenib binding mode within the CDK8 deep pocket in the CDK8/CycC complex provides valuable information for rational structure-kinetic-relationship (SKR) analysis of this important player in physiology and disease.

5. The discovery of slowness in terms of compound binding kinetics: a Structure-Kinetic-Relationship study on CDK8/CycC.

The high impact of crystallographic studies on the inhibitor's binding mode to its target has been recognized with pioneer studies such as bovine trypsin in complex with the bovine pancreatic trypsin inhibitor reconciling the kinetic data of inhibitor-enzyme interaction in the structural model (Blow et al., 1973; Rühlmann et al., 1973). The concept of structure activity relationship (SAR) describing the dependencies between binding affinity and compound structure was well explored in the last decades. However, similar studies describing the dependencies between binding kinetics and compound structure, the structure kinetics relationship (SKR) are missing.

5.1 The primary screen and the discovery of slowness

To find further compounds with an elongated residence time similar to the CDK8/CycC/sorafenib complex, a primary screen of parts of the Proteros library was performed against CDK8/CycC using the Proteros Reporter Displacement assay accordingly to the previously described method for active p38 α (Neumann et al., 2009; more detailed description in Neumann et al., 2011; chapter 2.15). In total 4921 compounds with different MW (fragment with MW <350g/mol and lead-like compounds with MW >350g/mol) were tested with hit rates as listed in Table 5.

	lead-like	hit-rate	fragment	hit-rate
tested compounds	2729		2192	
maximal assay concentration	1.91 μ M		38.2 μ M	
inhibition \geq 70%	68	2.5%	286	13.0%
inhibition \geq 80%	52	1.9%	201	9.2%
inhibition \geq 90%	64	1.8%	172	7.8%

Table 5 Results of the Primary Screen

210 compounds were selected for time-resolved IC₅₀ determination, which comprised fragments and lead-likes of different compound classes. As previously mentioned, compounds with long residence time values (equation 5, chapter 2.15) are characterized by slow binding kinetics which can be quantified by the value for the observed association rate k_{obs} ($k_{\text{obs}} = k_{\text{off}} (1 + [\text{cpd}]/K_d)$) (derived from

equation 2 and equation 4, chapter 2.15). In consequence compound target interactions with a long residence time are described by a low k_{off} and therefore also a low k_{obs} value for a compound concentration which is administered close to its K_d value. Since an IC_{50} measurement of the compound concentration covers the K_d value of the compound, compounds with slow binding kinetics are identified by the determination of the time dependent IC_{50} -values. Therefore a reporter probe, designed to bind to the active site in a “CDK8 DMG-in” conformation, and CDK8/CycC was mixed to form the reporter-target complex. In the following compounds were added at increasing concentration and IC_{50} -values were measured 15min and 80min after compound addition. Compounds with fast binding characteristics are expected to show the same IC_{50} -value after 15min and 80min. However, for slow binding compounds, that are characterized by a low k_{obs} value which goes along with a low k_{off} value and a long residence time, a decrease of the IC_{50} -value is expected with time. In figure 1a the corresponding IC_{50} -values were plotted with the value for 15min on the x-axis and the value for 80min on the y-axis. Compounds with fast binding kinetics are located on the bisecting line while compounds with slow binding kinetics are located below (Figure 34).

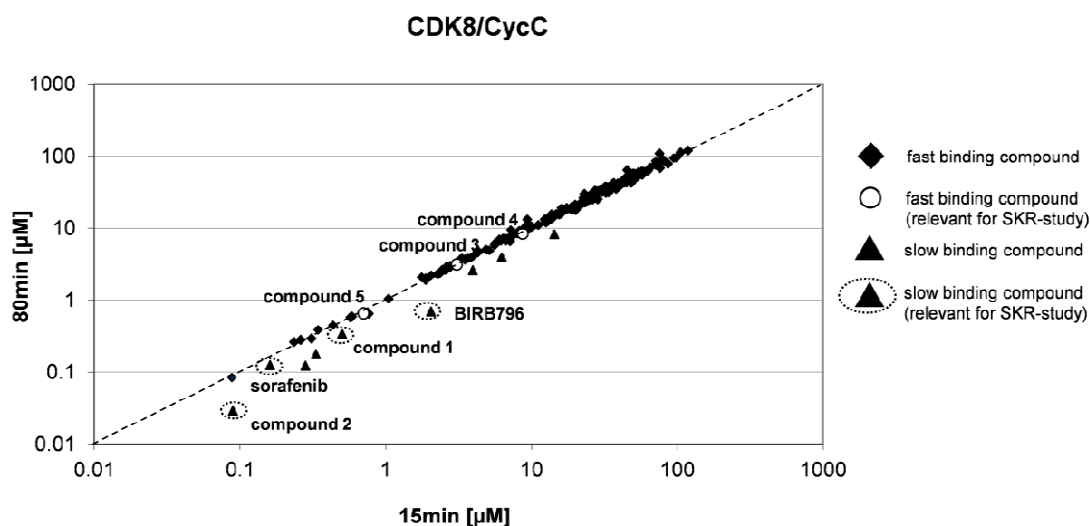


Figure 34 IC_{50} -determination of selected hit compounds of the primary screen

A slow binding behavior of compounds is revealed in a high throughput fashion even though only two time points (15min and 80min) are measured. Compounds that show the same IC_{50} value for both measurements have fast binding kinetics as compound binding has reached equilibrium quickly. They are located on the bisecting line. Compounds with different IC_{50} values for 15min and 80min have slow binding kinetics as compound binding is not completed after 15min. Those slow binding compounds are detectable below the bisecting line.

Even though binding kinetics is not measured continuously, the slow binding behaviour of nine lead-like compounds was easily detectable in this high throughput fashion (Figure 34). Amongst seven hits from different building blocks that showed slow binding kinetics, well-known slow binders such as sorafenib and BIRB796 are in accordance with previous reported results (Neumann et al., 2011).

Subsequently slow binding hit compounds were selected that share the same building block accordingly to common features of type II compounds. Compound 1 (N-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-4-[2-({[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]carbamoyl}amino)ethyl] piperazine-1-carboxamide) and compound 2 (*tert*-butyl[3-({[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]carbamoyl}amino)propyl]carbamate) consist of a central urea-core (Figure 35) that contacts the conserved K-E ion pair of all kinases involving the α C helix, the “roof” of the deep pocket (Simard et al., 2009b), and the backbone of the CDK8 DMG-motif. Thus a large hydrophobic moiety was focused that is also observed in BIRB796 (1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl], referred to as ‘deep pocket binding moiety’; Figure 35) to infiltrate the deep pocket. On the hinge-directed side, compound 1 and compound 2 vary with different substituents and functional groups since a hinge-interaction of the type II inhibitor is present in most but not all available kinase structures (Liu et al., 2006) and regarded as a favourable option to stabilize the protein-ligand interaction (Backes et al., 2010). The slow binding kinetics of compound 2 and compound 1 already allowed pre-anticipation of an extended residence time. In the next step it was interesting which compound structure similar to the building block of compound 2 and compound 1 would achieve a switch from fast to slow binding kinetics. This switch to emerging slow binding kinetics was possibly not monitored within the time interval of the first measurement point of IC₅₀-determination (Figure 34) after 15min and the subsequent time point after 80min. Consequently compound 2 and compound 1 building block-similar compounds were selected from the hit compounds (Figure 35; compound 3 (1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(2-hydroxyethyl)urea), compound 4 (1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-[2-(morpholine-4-yl)ethyl]-urea) and compound 5 (1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-[3-(morpholin-4-yl)propyl]urea)) and combined with newly designed compounds (Figure 35, compound 6 (1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]amine), compound 7 (1-[3-*tert*-

butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]urea), compound 8 (1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-methylurea), compound 9 (1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(3-hydroxypropyl)urea), compound 10 (1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(4-hydroxybutyl)urea) and compound 11 (1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(5-hydroxypentyl)urea)) to complete a back-to-front approach (Müller et al., 2010) for the SKR-study.

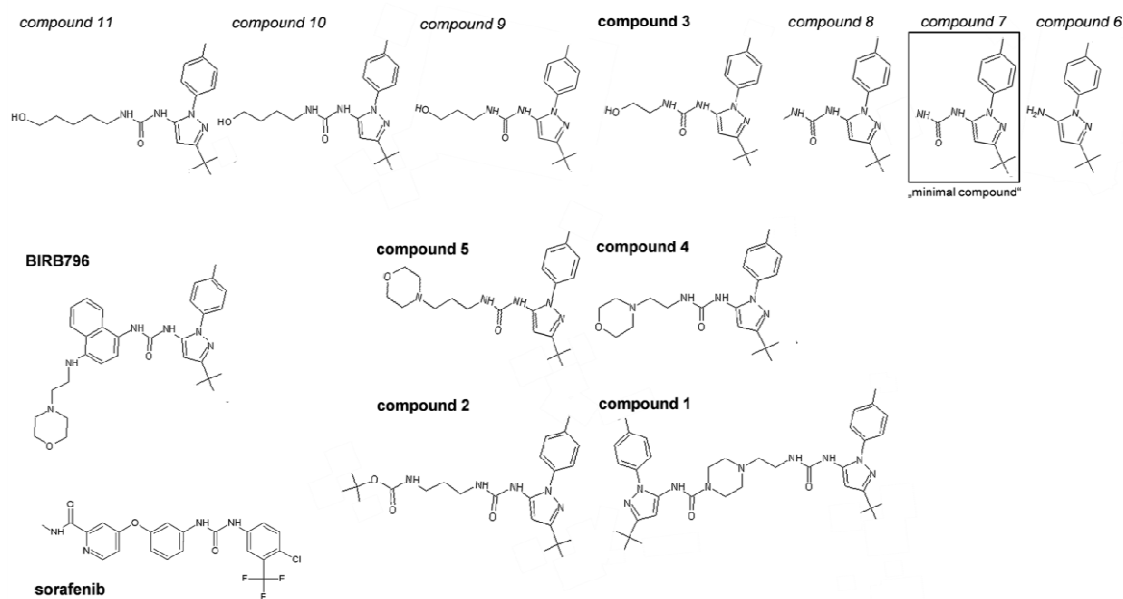


Figure 35 Structures of the compound series of the SKR-study with CDK8/CycC;

Hit compounds of the library screen are marked with bold letters; newly designed compounds to complete the back-to-front approach are marked with italic letters. **compound 1**: N-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-4-[2-({[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]carbamoyl}amino)ethyl] piperazine-1-carboxamide; **compound 2**: *tert*-butyl [3-({[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]carbamoyl}amino)propyl]carbamate; **compound 3**: 1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(2-hydroxyethyl)urea; **compound 4**: 1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-[2-(morpholine-4-yl)ethyl]-urea; **compound 5**: 1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-[3-(morpholine-4-yl)propyl]urea; **compound 6**: 1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]amine; **compound 7**: 1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]urea; **compound 8**: 1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-methylurea; **compound 9**: 1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(3-hydroxypropyl)urea; **compound 10**: 1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(4-hydroxybutyl)urea; **compound 11**: 1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(5-hydroxypentyl)urea;

Thus the 1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-moiety was supposed to serve as anchor-point within the deep pocket. This deep pocket binding moiety was combined with an amine-group (compound 6), a urea-group (compound 7) and a methyl-urea (compound 8) towards the fast binding hit compound ‘compound 3’ whose hinge-directed moiety consists of a (2-hydroxyethyl)-urea.

Within the hit compounds, the hinge-directed (2-hydroxy-ethyl)urea is substituted by a [2-(morpholine-4-yl)ethyl]urea (compound 4), that is further elongated to a [3-(morpholine-4-yl)propyl]urea (compound 5). Analogously the hinge-directed scaffold of the newly designed compounds varies with a (3-hydroxy-propyl)urea (compound 9), a (4-hydroxybutyl)urea (compound 10) and a 5-hydroxypentyl)urea (compound 11). All compounds were measured with the Reporter Displacement Assay continuously over 8h to allow very slow binding compounds to reach equilibrium (Table 6).

Compounds with fast binding kinetics for CDK8/CycC						
compound-ID		IC ₅₀ [μM]	K _d [μM]	k _{on} [s ⁻¹ μM ⁻¹]	k _{off} [s ⁻¹]	residence time [min]
compound 6	CDK8/CycC	>934.3	-	not applicable (n. a.)	n. a.	below detection limit (b. d.) < 1.4min
compound 7	CDK8/CycC	6.47	3.24	above detection limit (a. d.)	a. d.	b. d. < 1.4min
compound 8	CDK8/CycC	3.14	1.57	a. d.	a. d.	b. d. < 1.4min
compound 3	CDK8/CycC	11.6	5.82	a. d.	a. d.	b. d. < 1.4min
compound 4	CDK8/CycC	3.64	1.82	a. d.	a. d.	b. d. < 1.4min
compound 9	CDK8/CycC	7.06	3.53	a. d.	a. d.	b. d. < 1.4min
Compounds with slow binding kinetics for CDK8/CycC						
compound 10	CDK8/CycC	2.60	1.86	1.85×10 ⁻⁰³	2.41×10 ⁻⁰³	7
compound 5	CDK8/CycC	1.41	0.70	1.68×10 ⁻⁰³	1.18×10 ⁻⁰³	14
compound 11	CDK8/CycC	0.16	0.08	3.65×10 ⁻⁰³	2.90×10 ⁻⁰⁴	57
compound 2	CDK8/CycC	0.03	0.01	5.73×10 ⁻⁰⁴	8.57×10 ⁻⁰⁶	1944
compound 1	CDK8/CycC	0.06	0.06	2.99×10 ⁻⁰⁴	1.02×10 ⁻⁰⁵	1626

Table 6 Binding kinetics of the SKR-compounds

For a better overview the kinetic parameters of slow binding compounds are summarized on a kinetic map (Figure 36). Within the kinetic map, the k_{on} -value is plotted on the x-axis against the k_{off} -value on the y-axis. Thus the binding affinity ($K_d = k_{\text{off}}/k_{\text{on}}$) and the binding kinetics are visualized in parallel. Compounds with identical K_d values ($K_d = k_{\text{off}}/k_{\text{on}}$) are located on the same diagonal line. Compounds with high affinity are located in the lower right part of the kinetic map, while compounds with low affinity are located in the upper left part of the kinetic map. Compounds with longer residence time (residence time = $1/k_{\text{off}}$) are located below compounds with shorter residence times. The kinetic map allows to visualize if compound affinity improves due to increase of the k_{on} -rate or due to decrease of k_{off} -rate.

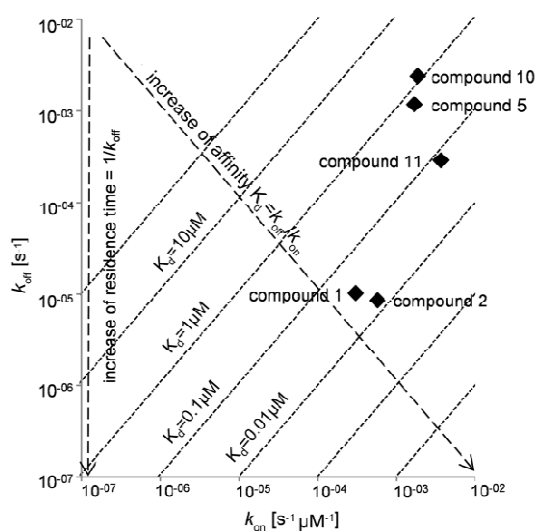


Figure 36 Kinetic map of the slow binding compounds

Binding affinity is visualized in parallel to the binding kinetics. Within the kinetic map, the k_{off} -rate is plotted on the y-axis and the k_{on} -rate is plotted on the x-axis. Thereby compounds with an identical K_d ($K_d = k_{\text{off}}/k_{\text{on}}$) lay on the same diagonal line. Arrows indicated that compounds with a low affinity are located in the upper left part of the kinetic map, while compounds with a higher affinity are located in the lower right part of the kinetic map. Compounds with a shorter residence time (residence time = $1/k_{\text{off}}$) are located above compounds with a longer residence time.

Interestingly, the deep pocket binding moiety in combination with an amine-extension (compound 6) did not show any detectable binding behavior (Table 6 and Figure 35). The ‘minimal compound’ of this inhibitor series exhibiting moderate affinity and a fast binding behavior for CDK8/CycC showed substitution of the amine-group by a urea-group (Table 6 and Figure 35 compound 7, the ‘minimal compound’). Extending the hinge-directed moiety of the minimal compound to an ethylurea in combination with different functional groups did not significantly change the compound affinity and

the fast binding kinetics (Table 6; compound 8, compound 3 and compound 4). Slow release of the compounds was observed only with compounds including a hinge-directed moiety that at least consists of a (4-hydroxybutyl)urea (Table 6; compound 10) or a (3-(morpholine-4-yl)propyl)urea (Table 6, compound 5). The residence time is extended analogously to increase hinge-directed moieties combining functional groups differing in their complexity (Table 6; compound 11, compound 2 and compound 1). In the case of compound 11, the k_{off} -rate ($2.9 \times 10^{-04} \text{ s}^{-1}$) is decreased by one order of magnitude compared to compound 10 ($2.41 \times 10^{-03} \text{ s}^{-1}$) while the k_{on} -rate is similar ($3.65 \times 10^{-03} \text{ s}^{-1} \mu\text{M}^{-1}$ and $1.85 \times 10^{-03} \text{ s}^{-1} \mu\text{M}^{-1}$). Thus the increase in affinity by a factor of 162.5 in compound 11 ($K_{\text{d}} = 0.0759 \mu\text{M}$) as compared to compound 10 ($K_{\text{d}} = 1.30 \mu\text{M}$) is obviously caused by the k_{off} -rate. This is also shown in the extension of residence time ($1/k_{\text{off}}$) of compound 11, that is increased 8-fold (57min) compared to compound 10 (7min). The illustration via a kinetic map enables direct selection of compounds with the required combination of affinity and residence time. Both lead-like compounds, compound 2 and compound 1, show a similarly high affinity (compound 2 with a $K_{\text{d}} = 0.015 \mu\text{M}$ and compound 1 with a $K_{\text{d}} = 0.03 \mu\text{M}$) with a favorable k_{off} -rate in compound 2 ($8.57 \times 10^{-06} \text{ s}^{-1}$) and compound 1 ($1.02 \times 10^{-05} \text{ s}^{-1}$) leading to an optimal residence time.

5.2 Combining structural with kinetic data reveals that the flip of the CDK8 DMG-motif does not influence the kinetic binding behavior of the compounds of the SKR series

In order to find structural parameters that trigger slow binding kinetics to CDK8/CycC, both fast and slow binding compounds were chosen as follows. Compound 7, the “minimal compound” compromises the basic scaffold of all selected compounds (1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]urea; referred to in the following as ‘minimal compound’). This minimal ensemble is identical to the deep pocket binding moiety of BIRB796 and its urea linker. The fast binding compounds compound 3 (Figure 35, (2-hydroxyethyl)-‘minimal compound’; $K_{\text{d}} = 5.82 \mu\text{M}$) and compound 4 (figure 2a, 2-morpholine-4-yl-ethyl-‘minimal compound’; $K_{\text{d}} = 1.82 \mu\text{M}$) differ within the functional groups of their hinge-directed moiety even though exhibiting comparably moderate affinity for CDK8/CycC. The compound 4 closely related compound 5 was chosen in order to structurally understand its flip from fast to emerging slow binding kinetics even though its K_{d} -value was only

increased by factor 3 (residence time: 14min; $K_d = 0.72\mu\text{M}$). Thereby the structure of compound 5 differs from compound 4 by the introduction of only one additional methyl-group within the linker of the hinge-directed morpholine-ring (Figure 35, 3-morpholine-4-yl-propyl-‘minimal compound’). The residence time of compound 11 (Figure 35, 4-hydroxypropyl-‘minimal compound’; $K_d = 0.16\mu\text{M}$) then is enhanced by a factor 7 (residence time: 57min) when elongating this hinge-directed extension to a 5-hydroxypentyl-urea (Figure 35; compound 11; 5-hydroxypentyl-‘minimal compound’). A clear step in residence time is achieved via the exchange of the morpholine-ring of compound 5 by a *tert*-butyl-propylcarbamate in compound 2 that results in an extreme shift of its residence time by factor 139 and its affinity by a factor 48 (residence time: 1944 min, $K_d = 0.015\mu\text{M}$; *tert*-butyl-propylcarbamate-‘minimal compound’). Within compound 1 the (2-morpholine-4-yl)ethyl-moiety of compound 4 is exchanged to a (2-piperazine-4-yl)ethyl-moiety in combination with an additional 1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]amide extending its residence time to 1626min and further increasing its affinity (residence time: 1626 minutes; $K_d = 0.034\mu\text{M}$; N-[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-4-[2-({[3-*tert*-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]carbamoyl} amino)ethyl] piperazine-1-carboxamide). The structures of these compounds within CDK8/CycC help to explain the observed kinetic data by showing how different groups interact in detail with CDK8/CycC, involving different parts and pockets of the CDK8 kinase. All compounds were back-soaked into CDK8/CycC crystals co-crystallized with a regular CDK8/CycC DMG-in type I compound to obtain the crystal structure with the CDK8/CycC complex (Table 7 and Table 8; Figure 37).

compound	7	3	4	5	11	2	1
	(4F6S)	(4F7J)	(4F70)	(4F6U)	(4F7N)	(4F7L)	(4F6W)
X-ray source	PXI/X06SA (SLS ¹)						
Detector	PILATUS 6M						
space group	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁
Cell							
a; b; c; [Å]	71.04; 71.39; 171.72	71.53; 71.91; 176.90	71.49; 71.85; 171.04	71.40; 71.02; 171.22	72.16; 71.76; 180.36	71.82; 71.24; 171.71	71.12; 71.71; 171.87
α ; β ; γ ; [°]	90.0; 90.0; 90.0	90.0; 90.0; 90.0	90.0; 90.0; 90.0	90.0; 90.0; 90.0	90.0; 90.0; 90.0	90.0; 90.0; 90.0	90.0; 90.0; 90.0
resolution[Å]	2.60 (2.81- 2.60)	2.60 (2.81- 2.60)	3.00 (3.24- 3.00)	2.10 (2.27- 2.10)	2.65 (2.86- 2.65)	2.90 (3.09- 2.90)	2.39 (2.58- 2.39)
unique reflections	27589 (5505)	28771 (5832)	18292 (3673)	51626 (10340)	27863 (5620)	19174 (2853)	35593 (7102)
multiplicity	8.1 (8.4)	5.4 (5.5)	8.0 (8.3)	8.1 (8.0)	3.7 (3.7)	3.9 (3.3)	8.1 (8.1)
completeness [%]	99.9 (99.7)	99.7 (99.4)	99.9 (99.7)	99.9 (100.0)	99.7 (99.6)	94.9 (84.3)	100.0 (100.0)
R _{sym} [%] ²	9.3 (63.5)	9.0 (50.4)	11.4 (51.1)	6.5 (54.9)	7.6 (51.5)	8.1 (43.2)	8.2 (59.4)
R _{meas} [%] ³	10.0 (67.7)	10.8 (61.2)	13.0 (58.2)	7.0 (58.7)	10.3 (68.7)	9.3 (50.9)	8.7 (63.5)
Mean(I)/sd ⁵	17.83 (4.54)	10.6 (2.5)	13.1 (4.1)	20.87 (4.43)	9.2 (2.0)	14.22 (3.29)	19.08 (4.30)

¹ SWISS LIGHT SOURCE (SLS, Villigen, Switzerland);

² values in parenthesis refer to the highest resolution bin;

³ calculated from independent reflections

Table 7 Data collection and processing statistics

compound	7 (4F6S)	3 (4F7J)	4 (4F70)	5 (4F6U)	11 (4F7N)	2 (4F7L)	1 (4F6W)
resolution [Å]	85.86- 2.60	88.45- 2.60	85.52- 3.00	85.61- 2.10	90.19- 2.65	85.85- 2.90	85.94- 2.39
number of reflections (working/test)	26652 / 936	27752 / 958	17610 / 630	49911 / 1715	26899 / 908	18512 / 662	34411 / 1181
R _{cryst} [%]	20.4	22.8	20.3	17.6	19.9	22.4	19.2
R _{free} [%] ²	26.2	27.6	23.9	19.8	22.8	27.0	23.4
amount of measured reflections within test set	3.4%	3.3%	3.5%	3.3%	3.3%	3.5%	3.3%
total number of atoms:							
protein	4925	4945	4919	4994	4993	4928	4951
water	90	58	32	309	61	25	150
ligand	20	23	28	29	26	31	47
deviation from ideal geometry:							
bond lengths [Å]	0.009	0.010	0.009	0.010	0.009	0.011	0.011
bond angles [°]	1.17	1.23	1.09	1.21	1.18	1.28	1.27
bonded B's [Å ²]	3.0	2.8	1.3	2.8	1.7	2.6	3.5
Ramachandran plot:							
most favoured regions [%]	93.8	92.1	92.7	93.3	94.6	90.1	94.4
additional allowed regions [%]	6.2	7.5	7.1	6.7	5.4	9.5	5.6
generously allowed regions [%]	0.0	0.4	0.2	0.0	0.0	0.4	0.0
disallowed regions [%]	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 8 Refinement statistics

Completeness of models defined by electron density is listed in the following. The final model compromises residues for CDK8 (Hs1-353) and CycC (Hs1-264) according to the sequence in Swiss-Prot ID P49336 and P24863. Additional residues were introduced (termed -1, 0, where 1 refers to M^{CDK8} and -3, -2, -1, 0 where 1 refers to M1^{CycC}) at the N-terminus of each protein during cloning at the cleavage site. compound 7: chain A (CDK8 residues 0-359) except for the segments 115-122, 178-194, 238-244 and 360-403; chain B (CycC residues -1-264) except for the segment 265-283; CDK8/CycC/compound 3: chain A (CDK8 residues 1-359) except for the segments 113-121, 177-194, 239-244 and 360-403; chain B (CycC residues -3-264) except for the segment 265-283; CDK8/CycC/compound 4: chain A (CDK8 residues 0-359) except for the segments 117-122, 175-194, 238-244 and 360-403; chain B (CycC residues -1-264) except for the segments 264-283; CDK8/CycC/compound 5: chain A (CDK8 residues -1-359) and chain B (CycC residues -2-264) except for the segments 116-120, 177-193, 240-244 and 360-403; CDK8/CycC/compound 11: chain A (CDK8 residues 0-361) except for the segments 115-120, 177-193, 240-244 and 362-403; chain B (CycC residues -2-265) except for the segment 266-283; CDK8/CycC/compound 2: chain A (CDK8 residues -1-359) except for the segments 115-121, 176-195, 239-244 and 360-403; chain B (CycC residues -1-264) except for the segment 265-283; CDK8/CycC/compound 1: chain A (CDK8 residues 1-359) except for the segments 116-120, 178-195, 239-244 and 360-403; chain B (CycC residues -2-264) except for the segment 265-283;

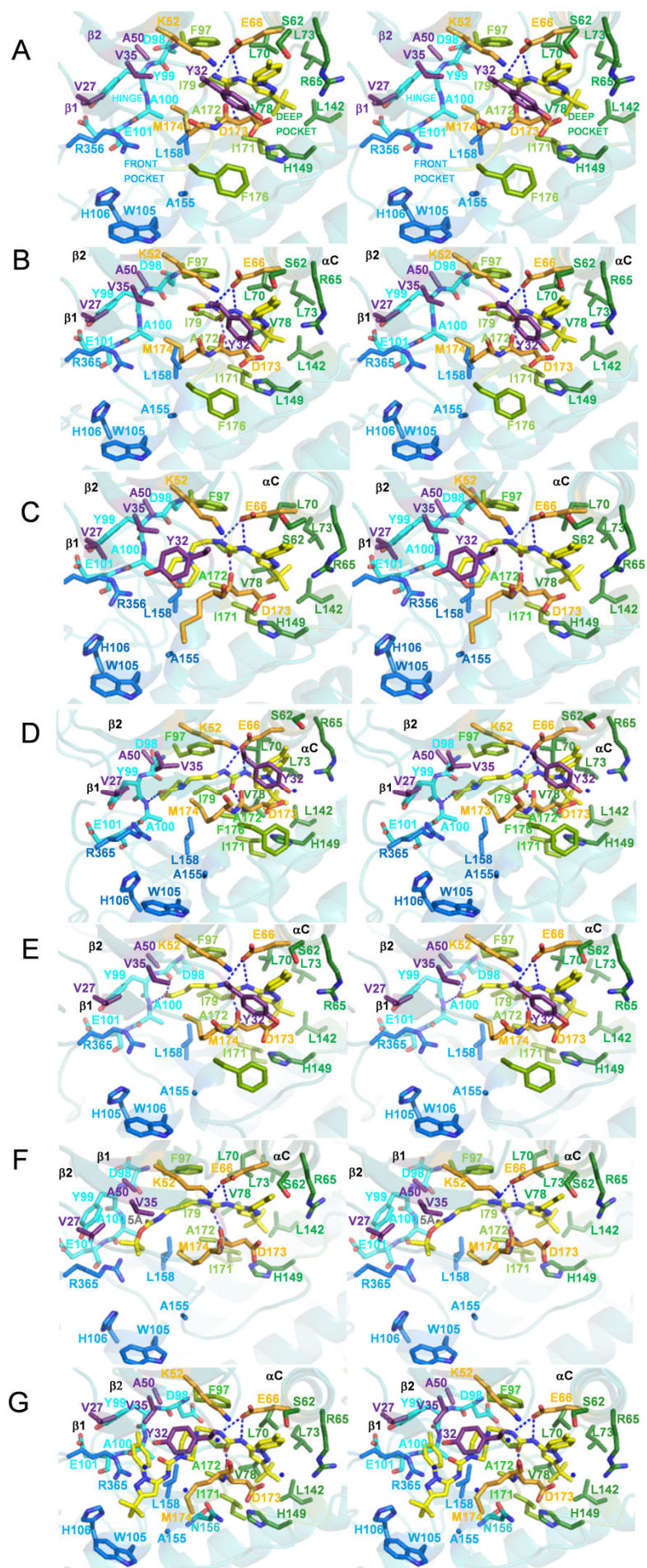


Figure 37 Detailed binding mode of the compounds of the SKR series (legend on next page)

Figure 37 Detailed binding mode of the compounds of the SKR series in complex with CDK8/CycC

The CDK8 residues are assigned to distinctive parts of the kinase (*teal*) and shown as sticks with the carbons colored as follows: the type II compound conserved interacting residues of the DMG-motif and the conserved kinase salt-bridge (*gold*), the DMG-motif surrounding residues (*slitpea green*), the deep pocket located and surrounding residues (*forest green*), the β 1- β 2 sheet located residues (*purple*), the hinge-region located residues (*turquoise*) and the front pocket located residues (*marine blue*). The carbons of the inhibitor are colored *yellow*. The other atoms are colored according to atom type: nitrogen (*blue*), oxygen (*red*), sulfur (*saffron*). In general hydrogen-bonds are indicated for compound-CDK8 interaction (*blue lines*). Water molecules that mediate hydrogen-bonding between CDK8 and the compound are shown as *blue* non-bound spheres but the lines indicating H-bonds are omitted for clarity. CycC is not shown for clarity as well. Secondary structural elements are indicated with black labels. All close-up stereo figures were positioned to allow best view depending on the inhibitor's binding mode. Slow binding of the compound is reconciled in the binding mode by either H-bonding with the hinge-region (D and E) or hydrophobic contacts within the front pocket (F and G). Further details see text. A) compound 7 complexed to CDK8/CycC with fast binding kinetics; interacting residues are indicated by labels in the previously defined colors the deep pocket (*forest green*), the hinge-region (*turquoise*) and the front pocket (*marine blue*); CDK8 activation segment not defined in electron density includes residues 178-194; B) compound 3 with fast binding kinetics residence time <1.4min; CDK8 activation segment not defined in electron density includes residues 177-194; C) compound 4 with fast binding kinetics residence time <1.4min; CDK8 activation segment not defined in electron density includes residues 175-194; D) compound 5 with emerging slow binding kinetics (residence time =14 min); CDK8 activation segment not defined in electron density includes residues 177-193; E) compound 11 with slow binding kinetics (residence time =57min); F) compound 2 with slow binding kinetics (residence time =1944 min; grey label indicates distance to hinge-region (5.1Å); Y32^{CDK8} is not shown as its side-chain was not completely defined by electron density; CDK8 activation segment not defined in electron density includes residues 176-195; G) compound 1 with slow binding kinetics (residence time = 1626min); additional hydrophilic interactions between compound 1 and the side-chain atoms of Y32^{CDK8} and N156^{CDK8} (*teal* with atoms colored according to type) are obviously bridged by water molecules. F176^{CDK8} is not shown as stick as only its main-chain was defined by electron density as well as the residues 177-195 of the activation segment.

Despite of its fast binding kinetics already the minimal compound, compound 7, (Figure 37A) induces a flip of the CDK8 DMG-motif to the out-conformation. Thereby 1-[3-tert-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-moiety (the ‘deep pocket binding moiety’) of compound 7 anchors within the CDK8 deep pocket and forms hydrophobic contacts with neighboring (Figure 37A, *forest green*) key residues of the DMG-motif and its surroundings (Figure 37A, *slitpea green*) and Y32^{CDK8} of the β 1- β 2 loop (Figure 37 3A, *purple*). The urea core of compound 7 contacts the essential E66^{CDK8} α C helix-residue (*gold*) and the D173^{CDK8} main-chain amide (*gold*) of the DMG-motif via hydrogen-bonding but does not disrupt the salt-bridge of the kinase-wide conserved E66^{CDK8}-K52^{CDK8} ion pair. This is probably resulting from the compound 7 structure that ends with the urea-group and cannot contact the CDK8 hinge-region (Figure 37A, *turquoise*) or its front pocket (Figure 37A, *marine blue*). The hinge-directed 2-hydroxyethyl-group of the fast binding compound 3 (Figure 37B) is also not long enough to contact the hinge-region so a standard interaction similar to the compound 7 (“the minimal compound”) binding mode to CDK8/CycC with few more contacts is observed. Even though the hinge-directed hydroxyl-group of compound 3 is located within hydrogen-bonding distance of K52^{CDK8}, H-bonding may be weak due to a disadvantageous angle. Analogously to the hinge-directed moiety of compound 3 (Figure 37B), the fast binding mode of compound 4 (Figure 37C) lacks a stabilizing interaction with the CDK8 hinge-region. While the hinge-directed 2-hydroxyethyl-group of compound 3 induces no significant further movement of CDK8 side-chains, the corresponding (2-morpholine-4-yl)ethyl-group of compound 4 obviously displaces the position of the Y32^{CDK8} side-chain towards the β 1-sheet. As a result the side-chain of M174^{CDK8} is moved even further than before, accompanied by disorder of G175-F176^{CDK8}. Thus the main-chain carbonyl of D173^{CDK8} appears in hydrogen-bonding distance of the nitrogen of the pyrazol-moiety but due to an unfavorable angle it rather forms an H-bond with E66^{CDK8} according to the overall conserved binding mode of the minimal compound 7 (Figure 37A). The slightly displaced binding mode of compound 4 appears to have only small impact on affinity ($K_d=1.82\mu\text{M}$) similar to the minimal compound 7 ($K_d=3.24\mu\text{M}$) and compound 3 ($K_d=5.82\mu\text{M}$) and their overall fast binding kinetics. More strikingly the lack of a hinge-directed fixation influences the velocity of the compound binding: the additionally introduced methyl-group in the hinge-directed moiety of compound 5 allows hinge-interaction by the (3-morpholine-4-yl) propyl-group (Figure 38D,

compound 5) with recognizable slow binding kinetics (residence time: 14 minutes). The conserved scaffold of compound 5 is similarly positioned as the minimal compound 7, and forms similar hydrophobic contacts and the standard hydrogen-bonds with CDK8/CycC as described for compound 7 and compound 3. The high resolution (2.1Å) of the X-ray structure of compound 5/CDK8/CycC allows the detection of a water-molecule that enables hydrogen-bonding between the nitrogen of the pyrazol-ring of the deep pocket binding moiety and R65^{CDK8}. The orientation of this arginine is similar for all structures and thus indicates the presence of a water molecule in all structures. The morpholine-group of compound 5 is anchored at the CDK8 hinge-region via H-bonding with the A100^{CDK8} backbone-amide (Figure 37D, turquoise). Interestingly, Y32^{CDK8} exhibits the same orientation as in the complexes of compound 7/CDK8/CycC (Figure 37A) and compound 3/CDK8/CycC (Figure 37B), which analogously allows interaction with the (4-methylphenyl)-1H-pyrazol-5-yl-moiety of the compound 5 deep pocket binding moiety. Furthermore interactions based on the hinge-directed morpholine-moiety of compound 5 with the β 1- β 2 loop and the DMG-motif surrounding residues increase the hydrophobic contact surface with CDK8/CycC compared to compound 4/CDK8/CycC. The residence time of compound 11 is obviously optimized by a factor 4 as its hinge-directed moiety, 5-hydroxypentyl-urea, forms a second hydrogen-bond with the CDK8/CycC region (D98^{CDK8} backbone-carbonyl and A100^{CDK8} backbone-amide). The conserved scaffold of compound 11 interacts with CDK8/CycC in a similar manner as described before. However, the extended residence time of compound 2 (Figure 37F, residence time 1944 minutes) is not explainable with an even tighter hinge-interaction. Whereas the conserved compound scaffold exhibits the standard binding mode of the minimal compound, the hinge-directed *tert*-butyl-propylcarbamate-moiety does not at all contact the CDK8 hinge-region (Figure 37F). Only the three standard hydrogen-bonds of the minimal compound are observable for compound 2: even though the main-chain amide of D173^{CDK8} is located in hydrogen-bonding distance of the nitrogens within the urea-linker of compound 2, the unfavorable angle does not allow for H-bond formation. The tertiary butyl of compound 2 is rather stabilized in a van der Waals distance of R356^{CDK8} that gets involved in this CDK8 front pocket interaction (Figure 37F *marine blue*). The effect of additionally targeting the CDK8 front pocket besides the deep pocket is most elaborated in the hinge-directed part of compound 1 (1-[3-*tert*-butyl-1-(4-methylphenyl)-1H-

pyrazol-5-yl]-amide-piperazine-ethyl- ‘minimal compound’, Figure 37G), which also does not interact with the CDK8 hinge-region. The piperazine-ring rather interacts with the Y32^{CDK8} side-chain that is flipped in a similar way as observed upon interaction with the corresponding 2-morpholine-4-yl-ethyl moiety in compound 4 (residence time <1.4 min) while the conserved scaffold of compound 1 still forms the standard contacts as observed for the minimal compound. Furthermore the interaction of the piperazine-moiety of compound 1 with the Y32^{CDK8} side-chain orientates the hinge-directed N-[3-tert-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]amide of compound 1 to extend further into the front pocket (Figure 37F, marine blue). Due to the resolution of this structure (2.4Å) additional water molecules involved in H-bonding between compound 1 and CDK8/CycC are observable within the CDK8 active site besides the water-molecule that mediates H-bonding with R65^{CDK8} within the CDK8 deep pocket similar to the compound 5/CDK8/CycC complex. Embracing the CDK8 front pocket an additional water molecule bridges a hydrophilic interaction between Y32^{CDK8} and compound 1. Similar to the compound 2/CDK8/CycC binding mode extensive contacts are formed between compound 1 and R365^{CDK8} (R365^{CDK8} π -stacking with the tolyl-moiety of compound 1) within the CDK8 front pocket including even more residues of the CDK8 front pocket directly (H106^{CDK8} and W105^{CDK8} of the α D helix located on the C-lobe) or indirectly involving water-molecules (N165^{CDK8} of the “ η 1”-₃₁₀helix/CDK8 C-lobe). However, despite of the increased interaction surface of compound 1 with CDK8/CycC the optimized residence time observed for compound 2 cannot be improved even more.

5.3 The X-ray structures of CDK8/CycC with slow and fast binding compounds indicate that the main contributions to residence time are conserved hydrogen-bonding and hydrophobic contacts within the CDK8 front pocket

Based on the structural analysis of CDK8/CycC in complex with both fast and slow binding diaryl urea compounds of this SKR series a working hypothesis was set up to identify parameters influencing the velocity of compound binding independent from the conformation of the T-loop and the DMG-motif. Within a SKR-study, the X-ray structures of the compound/CDK8/CycC complexes provide valuable information such as interactions within the binding surface for the optimization of the

residence time. Subsequently the contacts within the binding surface were analyzed per interacting atom of the corresponding residue in CDK8/CycC (Table 9).

crystal structure of CDK8/CycC							
complexed to compound	7	3	4	5	11	2	1
hydrogen-bonds	3	3	3	4	5	3	3
hydrophobic contacts	48 (30)	46 (33)	50 (30)	66 (48)	60 (41)	72 (46)	96 (66)
sum of contacts	51 (33)	49 (36)	53 (32)	70 (52)	63 (44)	78 (51)	99 (69)
contact atoms	29 (12)	30 (25)	32 (26)	47 (37)	37 (30)	43 (35)	58 (50)
residence time [min]	<1.4	<1.4	<1.4	14	57	1944	1626
quotient of residence time/ contact atoms	0.05 (0.11)	0.05 (0.06)	0.04 (0.05)	0.30 (0.4)	1.5 (1.9)	45.2 (55.5)	28.0 (32.5)

Table 9 Analysis of the compound binding surface per interacting atom CDK8/CycC

Hydrophobic contacts were calculated per atom (carbons, sulphur and halogens) in a van der Waals distance starting from 3.5Å up to 4Å, which is shown as number in brackets, allowing for deviations $\pm 0.2\text{\AA}$ to respect the rmsd and kinase dynamics of the crystal structure. Calculations were performed using CCP4, structure analysis, and manually corrected for chemical correctness. In the case of π -stacking we counted every atom of the respective arginine due to the electrochemical properties of this interaction. Hydrogen-bonding between the ligand and the inhibitor involving water-molecules were not calculated due to the differing resolutions of the crystal structures.

Besides the sum of hydrogen-bonds, hydrophobic contacts were calculated per atom (carbons, sulfur and halogens) using the sum of the Van der Waals radii + 0.5Å (contact atoms as defined by Lo Conte et al., 1999) that were increased by $\pm 0.2\text{\AA}$ in order to respect rmsd differences and kinase dynamics of the different CDK8/CycC crystal structures (resolution 2.1-3.0Å; calculations performed with CCP4, structure analysis). In the case of π -stacking every atom was counted in accord with the electrochemical properties of this interaction. H-bond formation involving water-molecules was not considered due to the different resolutions (2.1-3.0Å) of the X-ray structures as the presence of water molecules is uncertain at low resolution (Lo Conte et al., 1999). The sum of the binding interactions

within the compound-CDK8/CycC binding surface correlates with the extension of residence time and fast binding compounds form fewer hydrogen-bonds and hydrophobic contacts with CDK8/CycC (compound 7: 51; compound 3: 49; compound 4: 53; Table 9) than slow binding compounds (compound 5: 70; compound 11: 63; compound 2: 78; compound 1: 99; Table 9). The switch of a fast binding behavior to slow binding kinetics is best observable by the optimization of the binding mode of compound 5 whose hinge-directed linker to the morpholine-ring is elongated by one additional methyl-group with respect to compound 4. The slow binding kinetics of compound 5 is observed in parallel to an additional H-bond but also increased hydrophobic contacts (66 compared to 50). However, the binding mode of compound 11 to CDK8/CycC forming a second H-bond with the hinge-region compensates the fewer hydrophobic contacts with CDK8/CycC as compared to compound 5 and results in a fourfold optimized residence time. In general, type II compounds often form H-bonds to overall conserved kinase architecture such as the DFG-backbone, the K-E ion-bridge or the hinge-region. For the here presented SKR compound series, the key hydrogen-bonds of the compound's diaryl urea linker to the E-K ion-bridge and the DMG-backbone of both fast and slow binding compound are similar in the 'minimal compound' compound 7. Consequently the H-bond that compound 5 forms with the hinge-region is obviously triggering its slow binding behavior. Nevertheless, the addition of hydrogen-bonds with the hinge-region did only explain the appearance of slow binding kinetics for compound 5 forming one H-bond and improvement of residence time of compound 11 by two hydrogen-bonds. However, the extremely improved residence time of compound 1 and compound 2 obviously is independent of direct H-bonding of the compounds with the hinge-region or an additional H-bond besides the conserved ones. Further H-bonds via bridging water molecules are not considered for reasons given above, but may have a profound influence.

As a measure of the influence of the number of contacts on residence time the quotient of the residence time divided by the sum of each CDK8 atom interacting with the compound ("residence time per CDK8 atom", Table 9) was calculated. Fast binders start with a value around 0.05 increasing to 0.3 for emerging slow binding characteristic as in compound 5. The extended residence time as observed for compound 11 shows a quotient of 1.5 that is increased to a value of 28 (compound 1; 58 CDK8 atoms) and 45.2 (compound 2; 43 CDK8 atoms) for these most optimized compounds.

Apparently not only the sum of the interacting atoms in CDK8 (as observed for compound 1) but also their properties and location within the kinase have an influence on efficient binding of a compound. To explain this behavior the different binding modes of the SKR compounds by were reconsidered by quantifying hydrophobic contacts of the compounds (Figure 38): thus a site-specific interaction is related to an optimization of the residence time.

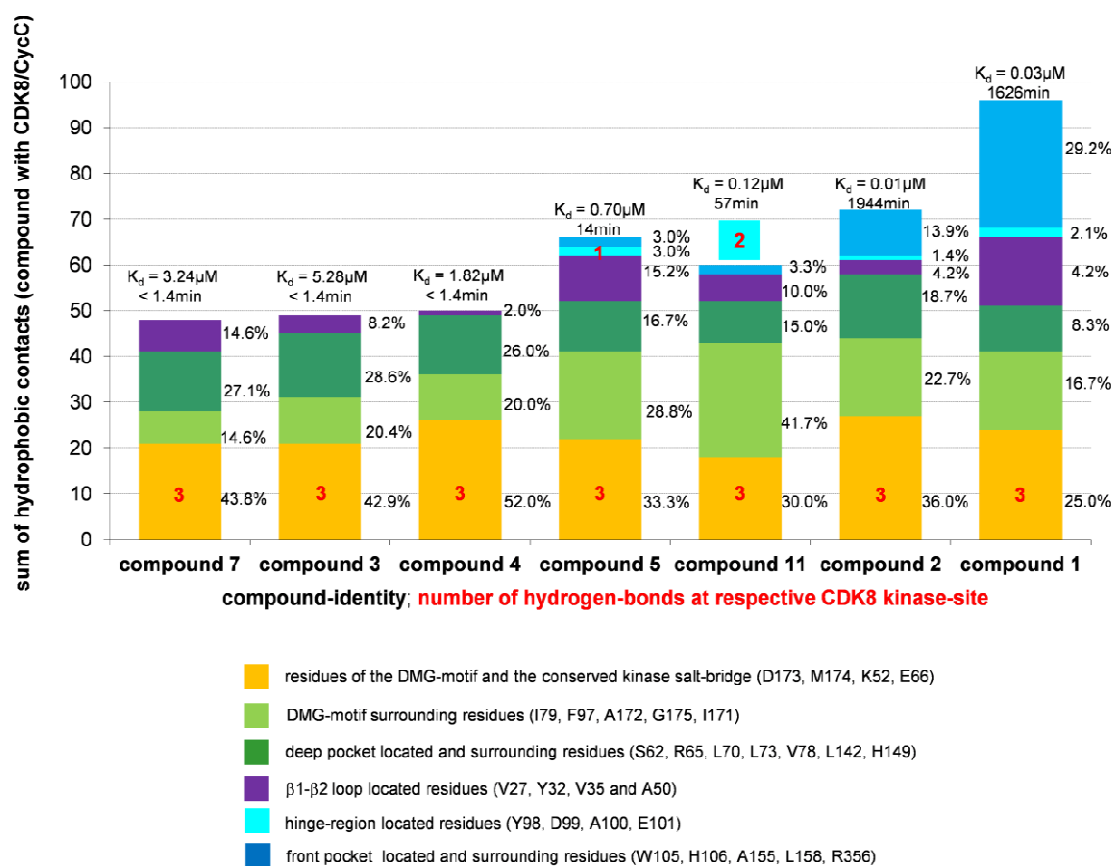


Figure 38 Overview of the fraction of the respective aminoacids on the sum of hydrophobic interactions calculated per atom

In this plot, the sum of the hydrophobic contacts per atom is plotted on the y-axis for the corresponding X-ray structures of the CDK8/CycC/compound complexes. Labels indicate the K_d -value [μM] and residence time [min] of the corresponding CDK8/CycC/compound complex. The number of H-bonds between compound and at the respective interacting kinase part is indicated with red labels. As compound 11 forms no hydrophobic contact with the hinge-region but 2 hydrogen-bonds the red label is underplayed with a turquoise color. The hydrophobic contacts were calculated using CCP4 structure analysis within a distance of $3.5\text{-}4.0\text{\AA} \pm 0.2$ to consider the rmsd and kinase dynamics of the respective X-ray structures. In the case of π -stacking, every atom was counted to respect the electrochemical properties of this interaction. Thereby CDK8 residues were assigned to parts of the kinase as follows: the residues of the DMG-motif and the conserved kinase salt-bridge (*gold*), the DMG-motif surrounding residues (*slitpea green*), the deep pocket located and surrounding residues (*forest green*), the β 1- β 2 sheet located residues (*purple*), the hinge-region located residues (*turquoise*) and the front pocket located residues (*marine blue*). The number of hydrogen-bonds is indicated on the respective kinase region (*red*).

For the fast binding compounds (compound 7, compound 3 and compound 4) the major part of hydrophobic interaction is based on the key residues of the DMG-motif and the kinase-wide conserved E-K ion bridge (43.8%, 42.9% and 52.0% of total; Figure 38) and explains their ability to change the conformation of the DMG-motif. The fast binders lack interactions with the hinge-region and form only few contacts around the edge of the front pocket (L158^{CDK8}). In sum the constant fraction of hydrophobic interactions within the deep pocket region and a similar fraction of hydrophobic contacts involving the DMG-motif surrounding residues explain the binding mode of these compounds. The fraction of hydrophobic contacts with the β 1- β 2 loop located residues differs due to the varying conformation of Y32^{CDK8} (Figure 37A-C; Figure 38). However, the relative importance of these “key interactions” is reduced in accordance to the binding mode and structure of the compound when additional parts of the kinase are targeted by hydrophobic interactions.

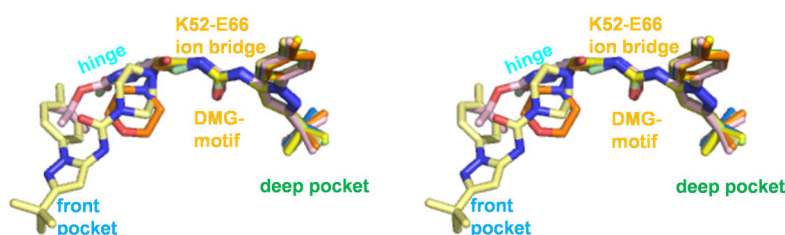
Similar to the analysis of H-bond formation, slow binding kinetics is only detectable for compounds that show clearly increased numbers of hydrophobic interaction in the β 1- β 2 region together with the DMG-surrounding motif and the hinge-region (compound 5, compound 11, compound 2 and compound 1; Figure 38). Interestingly the optimization of residence time for compound 11 compared to compound 5 was obviously obtained via a second hydrogen-bond with the hinge-region but no further hydrophobic contacts, which highlights the impact of hydrogen-bonding with the hinge-region. However, the increased part of hydrophobic interaction with the CDK8 front pocket located residues as observed for the compounds, compound 2 (13.9% of total; residence time: 1944min; Figure 38) and compound 1 (29.2% of total; 1626min; Figure 38) seems to optimize their residence time most strikingly, even though no direct H-bonding of these compounds with the hinge-region is observable and further hydrophobic contacts do not significantly contribute to the interaction (compound 1: 2.1% of total and compound 2 1.4% of total Figure 38). Even though it appears that the sum of contacts correlates with a long residence time as obvious from SKR studies on compound 1, these contacts should be fine-tuned involving less CDK8 atoms that more efficiently allow optimizing the residence time as observed for compound 2 (Table 9).

In sum the analysis of the compound-CDK8/CycC interactions reveals that three standard hydrogen-bonds and hydrophobic interaction of the SKR compound series as already observable for its minimal compound are the pre-condition for a compound to flip the DMG-motif from CDK8 DMG-in to DMG-out. Both polar and non-polar contacts of the compounds with the hinge-region are able to evoke slow binding kinetics. Additional hydrogen-bonding with the hinge-region appears to be indispensable to improve the k_{off} -rate. The extension of residence time of a compound in the here presented set-ups was triggered by additional hydrophobic contacts within the CDK8 front pocket, which additionally involves the R356^{CDK8} (C-terminal segment of CDK8) and N156^{CDK8} (C-lobe). According to these observations of the here presented SKR-study the emerging slow binding kinetics of compound 10 (7min) can be explained even though of no structure of compound 10 complexed to CDK8/CycC is available. Due to its compound structure, the hinge-directed 4-hydroxybutyl of compound 10 should be able to form one hydrogen-bond with the hinge-region similar to the elongated 5-hydroxypentyl of compound 11. However, the extension of compound 10 is expected to be too short to form a second H-bond with the hinge-region. As compound 10 is supposed to establish one hydrogen-bond with the hinge-region similar to the binding mode of compound 5, its similarly emerging slow binding kinetics (residence time: 7min compared to 14min in compound 5) is explainable from its structure. The lack of a second H-bond with the hinge-region as observed for compound 11 or any further hydrophobic contacts within the front pocket should be sufficient to explain the factor 8 decreased residence time of compound 10 compared to compound 11.

5.4 Discussion

The here presented crystal structures of CDK8/CycC in complex with both fast and slow binding compounds that address the deep pocket of a kinase reveal that other factors than the change of the conformation of the CDK8 DMG-motif increase the residence time of the CDK8/CycC specific compounds. Both fast and slow binding compounds of the CDK8/CycC SKR series enter the deep pocket and induce a CDK8 DMG-out conformation in the CDK8/CycC complex. The positions of the deep pocket binding moiety and the urea-linker of all compounds align closely in an overlay (Figure 39), and obey a standard binding mode of the compound scaffold as observed for the minimal compound (compound 7).

A overlay of all SKR compounds



B compounds turned by 45 degree around x-axis

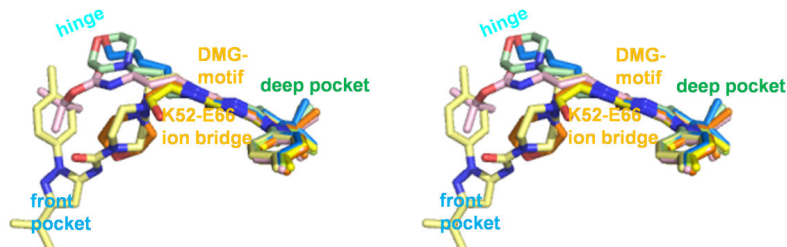


Figure 39 Superposition of the compounds of the SKR series

The overlay of the compounds is centered on CDK8 of the CDK8/CycC/compound 7 (4F6S) X-ray structure comprising the basic scaffold of the compound structure within the “minimal compound” (CDK8/CycC/compound 3 (4F7J): rmsd 0.4Å; CDK8/CycC/compound 4 (4F70): rmsd 0.3Å; CDK8/CycC/compound 5 (4F6U): rmsd 0.4Å, CDK8/CycC/compound 2 (4F7L): rmsd: 0.4Å; CDK8/CycC/compound 1 (4F6W): rmsd 0.4Å; CDK8/CycC/compound 11 (4F7N): rmsd 0.3Å; all X-ray structures share a sequence identity of 100%). compound 7 (carbon: *limon*), compound 3 (carbon: *yellow*), compound 4 (carbon: *orange*), compound 5 (carbon: *pale green*), compound 11 (carbon: *marine blue*), compound 2 (carbon: *rose*) and compound 1 (carbon: *pale yellow*). For clarity, CDK8/CycC residues were omitted and only important kinase-parts are indicated with labels. Even though superposition was centered on CDK8, both the central urea linker and the deep pocket binding moieties of all compounds can be aligned closely. On contrary the orientation and position of their hinge-directed functional groups differs.

Neither the degree of disorder of the CDK8 T-loop residues nor its displacement seems to influence the binding kinetics of the compounds. The side-chain of D173 (DMG-motif) is found in similar positions for both fast and slow binding compounds. Only the side-chain of M174 (DMG-motif) is varying according to the location of the Y32^{CDK8} side-chain (β 1- β 2 loop) in the compound 4/CDK8/CycC and compound 1/CDK8/CycC complex which allows interaction of the side-chain of Y32^{CDK8} with the piperazine-moiety of compound 1. Interestingly Y32^{CDK8} corresponds to the kinase-wide conserved aromatic residue which caps the site of the phosphate transfer during ATP binding (Bossemeyer et al., 1993; Huse et al., 2002). For CDK2 the corresponding Y52^{CDK2} was reported to distort the position of ATP binding and the catalytic step of phospho-transfer when phosphorylated *in-silico* (Bartova et al., 2004) and thus high flexibility of this side-chain was expected. The conserved scaffold of the compound structure of this SKR series (Figure 39) exhibits a standard binding mode of the diaryl urea-core and the deep pocket binding moiety, which is generally reported for DFG-out binders such as BIRB796 (Dietrich et al., 2010; Backes et al., 2008). But only H-bonding with the hinge-region triggers slow binding behavior of the SKR compounds that is optimized by an increased hydrophobic interaction with the CDK8 front pocket. This is consistent with literature reporting that H-bonding of a compound with a more solvent-shielded slows down its dissociation rate (Schmidtke et al., 2011). However, the observation that fast binding kinetics of the compounds of this SKR series induce a DMG-out binding mode is contrary to literature, as it was shown previously for the p38 α mitogen activated protein kinase (MAPK) that compounds belonging to the diaryl urea class or BIRB796 induce a DFG-out conformation with associated characteristic slow binding kinetics (Pargellis et al., 2002, Regan et al., 2003). The slow binding kinetic signature is believed to result from the compound selecting the inactive T-loop kinase conformation from a preexisting equilibrium (Backes et al., 2008). In fact CDK8 is expected to preferentially adopt the DMG-in kinase conformation at least when complexed to CycC. The deep pocket binding moiety of this series in combination with an amine-group (compound 6) does not bind to CDK8/CycC which indicates that the deep pocket in CDK8 is not accessible for this compound. The minimal compound to open the deep pocket contains at least an additional urea linker despite of an amid-group. Therefore the H-bonding between the urea-linker and the α C helix/DMG-backbone is regarded as “key interaction” to

provide sufficient binding energy together with the compounds' deep pocket binding moiety to enable the flip of the DMG-motif from an in- to an out-conformation. The kinetic results of the Reporter Displacement assay that slow binding compounds displace a DMG-in conformation binding probe are justified in the X-ray structure of CDK8/CycC in complex with the compounds of the here presented SKR series. Indeed, the binding of the compounds changes the conformation of the CDK8 kinase to the DMG-out conformation even within the CDK8/CycC/type I compound from the DMG-in conformation crystal lattice. However, in the case of CDK8/CycC, this exception to the expected slow type II binding mode might be attributed to the interaction of CDK8 with CycC. Previous to compound binding the association of the cyclin to the α C helix should already account for the selection and stabilization of the open CDK8 conformation within the CDK8/CycC complex. Within an isolated CDK8 the α C helix should be rotated outside the CDK binding site as it lacks the interaction with CycC thereby positioning E66^{CDK8} (α C helix) at least unfavorable for binding with a type II compound.

Besides the H-bonds with the hinge-region the quotient of residence time divided per the sum of interacting atoms of CDK8 was introduced to quantify the optimization of the CDK8/CycC-compound interaction surface as an additional metric. Similarly metrics were described for binding kinetics of a compound without the respective structure to rank the physiochemical properties of a compound series with similar MW and residence time, the "kinetic efficiency" (KE), as the kinetic parameters for the binding interaction per atom of a compound ($KE = \tau / N_{HA} = t_{1/2} / (0.693 \times N_{HA})$; where τ is the residence time and N_{HA} is the number of non-hydrogen atoms of a compound). Even though, the sum of interaction needs to be fine-tuned reconsidering the distinct location of the compound-interacting residues within CDK8/CycC. Especially targeting the CDK8 front pocket where the C-terminal segment of the CDK8-construct (R356^{CDK8}) gets involved optimizes the residence time for CDK8/CycC compounds. A kinase-site dependent increase of affinity was as well reported within the CDK-family for type I compounds that bind the CDK front pocket, the so-called "Lys89-pocket" (reviewed by Lolli, 2010). Moreover the C-terminus of CDK9 within the CDK9/CycT complex was

reported to interact with the CDK9 hinge-region and contribute to an increased compound affinity (Baumli et al., 2011).

The impact on the extension of the residence time by hydrophobic interactions within the hinge-region/front pocket appears explainable as well from theories on kinase dynamics such as the ‘hydrophobic spine’ kinase theory (Kornev et al., 2006, Müller et al., 2010). The ‘hydrophobic spine’ contains four amino acids (L76, V78, H149 and M173 in CDK8 by similarity to PKA) that are considered to build a hydrophobic network/spine allowing the kinase “breathing motion” between the N-lobe and the C-lobe to enable nucleotide binding and release (Noble et al., 2005). In the case of the CDK/cyclin family this mechanism is modified as exemplified in CDK2 binding of CycA, which positions the CDK loop accurately for the active kinase conformation. However, the disruption of the CDK8 T-loop/CycC interaction by a compound without additional fixation within the hinge-region or front pocket seems not sufficient to inhibit a “breathing motion” that releases the compound. This possibly results in a fast dissociation rate of the compound with fast binding kinetics. Only in compounds additionally interacting with the hinge-region of CDK8 or even more efficiently with the CDK8 front pocket as “anchor point” the disruption of this CDK8 T-loop/CycC interaction surface seems to be stabilized. Thereby the possibly more efficient disruption of the kinase motions seems to explain the slower and inhibited dissociation of the compound/ compound release. Analogously dynamics in terms of internal motions of free CDK2 were reported to be changed after binding of inhibitors to a greater extent than did the static protein conformation (Otypeka et al., 2002). X-ray crystallography as a static model limits our possibilities to detect a hindered kinase motion. However, the combination of the data on the X-ray structures with kinetic data upon the respective compound binding within the here presented SKR study might be interpreted that slow binding kinetics may be driven by a disabled kinase motion. More extended SKR studies on other targets as demonstrated here on the example of CDK8/CycC might be a useful tool to find further clues on parameters that influence the binding kinetics of a compound.

6. CDK8/CycC in the DMG-in kinase conformation versus the CDK8/CycC DMG-out kinase conformation: the two faces of the CDK8/CycC complex

As no activatory phosphorylation of the CDK8 T-loop has been observed yet in contrast to the general mechanism of CDK activation (Jeffrey et al., 1995; Pavletich, 1999), a peculiarity of the CDK8/CycC kinase activation was preanticipated. Based on the comprehensive structural information on CDK2 structures it was discovered that the general CDK activation process requires two steps including binding of the cyclin to rearrange the CDK α C helix towards the activation segment and followed by reversible phosphorylation of the CDK T-loop: Subsequently an electrostatic network with a conserved arginine triade interacts with the phospho-residue to maintain the activation segment in an open conformation. However, several crystal structures have already questioned the universal validity of this two step activation model such as CDK4/CycD (Day et al., 2009; Takaki et al., 2009).

6.1 Binding assay to select potential CDK8/CycC inhibitors that target the CDK8 DMG-in kinase conformation

In order to find suitable compounds to bind a CDK8/CycC DMG-in conformation, the previously performed screen of parts of the Proteros Fragments library for the SKR-study (see chapter 5) was reconsidered. Several fast binding compounds had been detected (Figure 40), that are expected to show a constant IC_{50} value over time. These are located on the bisecting line of a scatter plot with the respective IC_{50} -values after 15min plotted on the x-axis and the IC_{50} values after 80min plotted on the y-axis (see chapter 5).

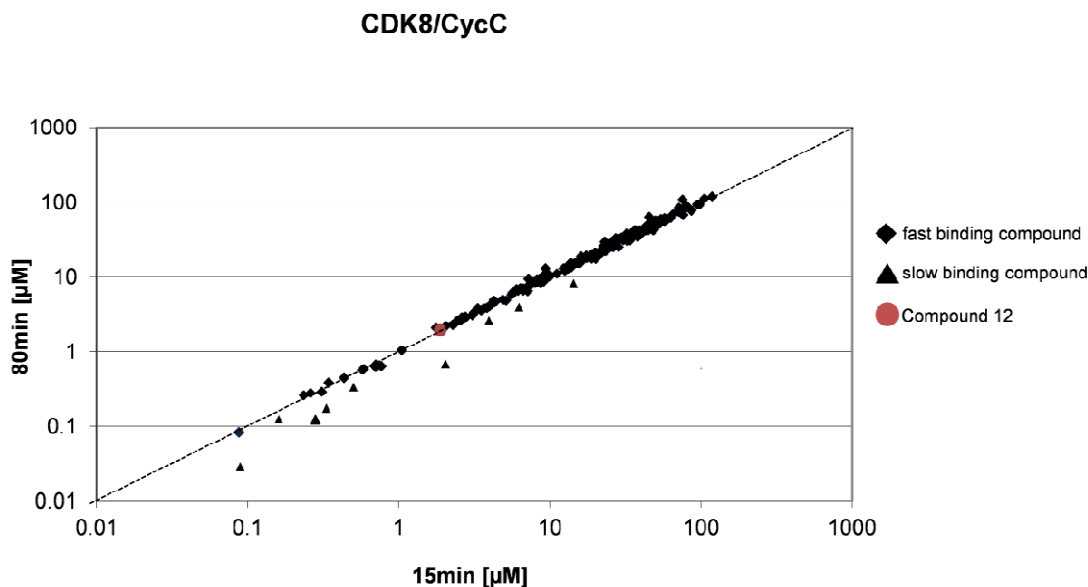


Figure 40 Evaluation of the library screen of parts of the Proteros library for fast binding compounds

IC₅₀-determination of selected hit compounds of the primary screen measured at two time points (15min and 80min). Compounds that show the same IC₅₀ value for both measurements have fast binding kinetics as compound binding has reached equilibrium quickly. They are located on the bisecting line. Compounds with different IC₅₀ values for 15min and 80min have slow binding kinetics as compound binding is not completed after 15min. Those slow binding compounds are detectable below the bisecting line.

Several compounds of these fast binding compounds were evaluated for crystallization with CDK8/CycC. Best results were obtained from crystals of CDK8/CycC in complex with N-(2-phenylethyl)-quinazoline-4-amine (compound 12) due to its compound structure and expected solubility (Table 10, 4F7S).

Data Collection	
Beamline	SLS X06SA
Space group	P 2 ₁ 2 ₁ 2 ₁
	a= 70.50Å; b= 70.77Å; c= 170.14Å
Unit cell dimensions	$\alpha= 90^\circ$, $\beta= 90^\circ$, $\gamma= 90^\circ$
Data Processing	
Resolution (Å)	85.07-2.20
R _{sym}	7.4 [68.3] ^a
Total number of observations	238251 [43856]
Total number unique observations	43856 [7387]
Mean I/sd(I)	15.60 [2.76]
Completeness	99.5 [99.5]
Multiplicity	5.4 [5.6]
Refinement	
Total number residues ^b	543
Number of water molecules	209
R _{Cryst}	19.3
R _{Free} ^c	22.2
r.m.s. bonds (Å)	0.01
r.m.s. angles (deg.)	1.18
Rachmandran plot	
Most favoured geometry (%)	94.3
Additionally allowed (%)	5.7

Table 10 Crystallographic parameters for human CDK8/CycC/compound 12 (4F7S)

^a Values in brackets refer to the highest resolution shell; ^b The final model comprises residues the CDK8 residues 1-359 and CycC residues 1- 264 with three additional residues at the N-terminus of CycC from expression tag (referred to as D-2 D-1 and A0). Segments that are not defined in electron density include the segments 116-121, 187-195, 238-242 and 360-403 for CDK8 and the segment 265-283 for CycC. ^c Test-set contains 3.3 % of measured reflections.

In presence of compound 12 within the active site of the CDK8/CycC/compound 12 crystal structure (4F7S) reveals a CDK8 DMG-in kinase conformation. Comparison of this overall CDK8/CycC/compound 12 structure with the previously solved crystal structure of CDK8/CycC/sorafenib (3RGF) in the CDK8 DMG-out kinase conformation several previously disordered are now well-defined (Figure 41). These include the CDK8 β 1-sheet, a segment of the

CDK8 C-terminus and parts of the activation segment. The CycC structure is almost identical in both CDK8/CycC structures. However, the CycC H3-H4 loop is shifted towards CDK8 in the active CDK8/CycC DMG-in kinase conformation. This observation may be interpreted as part of an activatory interaction based on CycC within the CDK8 DMG-in kinase conformation, which will be described in detail later on.

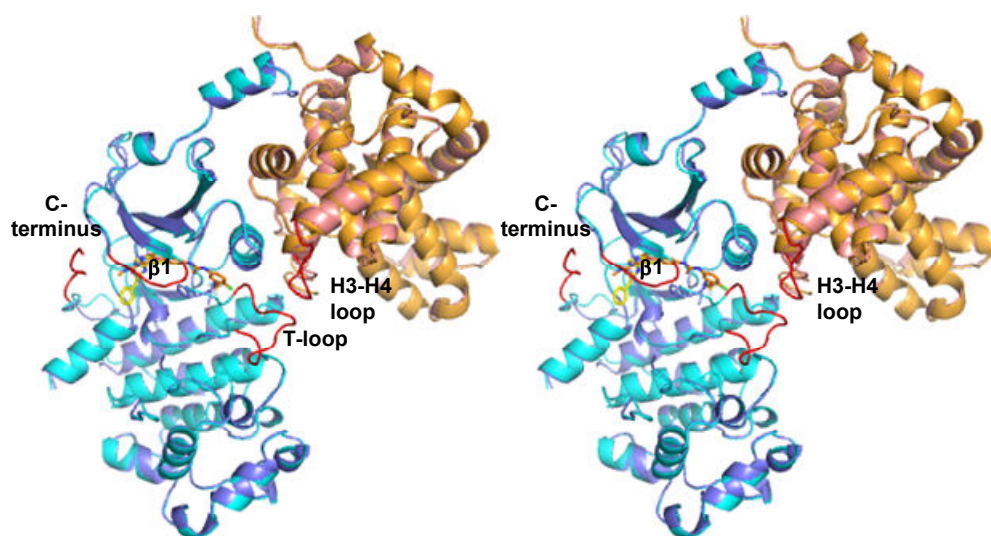


Figure 41 CDK8/CycC/compound 12 (4F7S) superimposed to CDK8/CycC/sorafenib

3RGF (CDK8^{3RGF}: slate blue; CycC^{3RGF}: salmon; sorafenib: orange); 4F7S (CDK8: turquoise; CycC: gold; compound 12: yellow); atoms colored according to type with nitrogen blue; oxygen red, chloride green and fluorine green; core rmsd: 0.56Å; differing areas are colored red and labeled.

6.2 Structural data on compound 12 in the active site CDK8/CycC site

A close-up of the activation segment of the new CDK8/CycC/compound 12 structure reveals that the kinase-wide conserved salt-bridge involved in ATP-binding (Bossemeyer et al., 1993; Huse et al., 2002) is established involving K52^{CDK8} and E66^{CDK8}. The side-chains of M174^{CDK8} and F176^{CDK8} are located within the CDK8 deep pocket while the side-chain of D173^{CDK8} exhibits a solvent-exposed conformation. Based on M174^{CDK8} and F176^{CDK8} van der Waals contacts are formed with residues located nearby the catalytic cleft (L69^{CDK8}, L73^{CDK8}, V147^{CDK8}, and E66^{CDK8}), which were shown

previously to be targeted by deep pocket binders within the CDK8/CycC DMG-out conformation (see chapter 5). Compound 12 is complexed (Figure 42) at the hinge-region of the catalytic cleft between the two lobes of CDK8 at the nucleotide binding site located below the $\beta 1$ - $\beta 2$ loop in a ATP-competitive type I inhibitor binding manner (Bossemeyer et al., 1993; Huse et al., 2002).

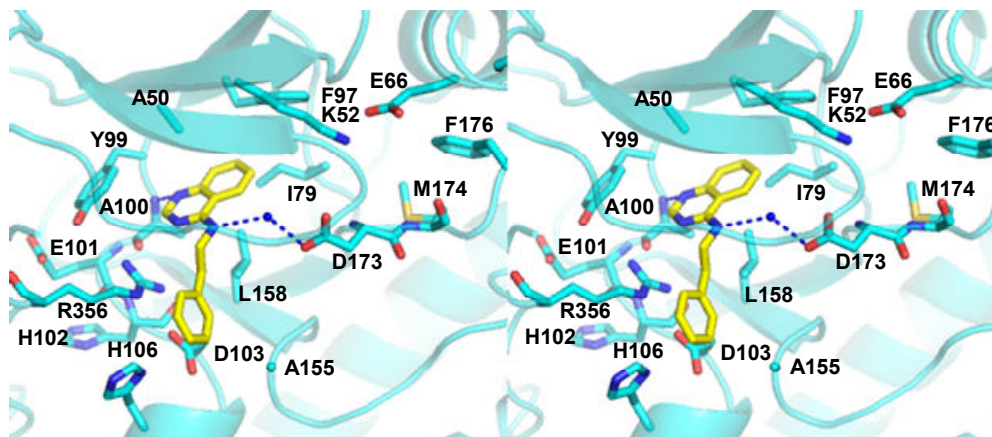


Figure 42 The ATP competitive type I binding mode of compound 12 targets CDK8/CycC within the DMG-in kinase conformation

Close up of the active site of CDK8 (turquoise); residues colored according to atom type (oxygen red, nitrogen blue, sulphur saffron); compound 12 (carbon yellow; nitrogen blue); interacting residues within a distance of compound 12 $\leq 4\text{\AA}$ are shown as well as neighboring residues of the DMG-motif and the conserved K-E ion bridge.

The quinazoline-moiety of compound 12 establishes a hydrogen-bond with the hinge-region ($A100^{\text{CDK8}}$ main-chain carbonyl) and several hydrophobic contacts with neighboring pocket forming residues ($A50^{\text{CDK8}}$, $I79^{\text{CDK8}}$, $F97^{\text{CDK8}}$ and $Y99^{\text{CDK8}}$). Furthermore the nitrogen within the quinazoline-4-amine-moiety contacts the side-chain of $D173^{\text{CDK8}}$ via hydrogen-bonding involving a water molecule. Interestingly the side-chain of $R356^{\text{CDK8}}$, located within the C-terminal CDK8 construct, exerts hydrophobic contacts with the inhibitor. Moreover $R356^{\text{CDK8}}$ establishes hydrophilic contacts with residues of the hinge-region involving H-bonding between the $R356^{\text{CDK8}}$ side-chain with the $A100^{\text{CDK8}}$ main-chain carbonyl, the $H102^{\text{CDK8}}$ main-chain carbonyl and the $D103^{\text{CDK8}}$ side-chain (Figure 42) whereas the $R356^{\text{CDK8}}$ main-chain amide contacts the $E101^{\text{CDK8}}$ side-chain. In the CDK9/CycT1 complex, a similar arrangement of the C-terminal part of the CDK9 construct towards the hinge-region was reported to contribute to inhibitor affinity (Baumli et al., 2010). The same effect could explain the comparatively moderate affinity of compound 12 despite of its sparse interaction surface with CDK8/CycC. The phenyl-moiety of compound 12 exerts hydrophobic contacts with residues forming

the CDK8 front pocket (H106^{CDK8}, Asp103^{CDK8}, Leu158^{CDK8} and A155^{CDK8}). The same area is described as the “Lys89 pocket” for CDKs (based on the H106^{CDK8} corresponding residue Lys89^{CDK2}) where compound binding results in increased selectivity as exemplarily shown for compounds with fine-tuned electrostatic contacts for CDK2 or hydrophobic interaction with CDK4 and CDK6 (Lolli, 2010).

6.3 Arrangements of CDK8/CycC related to the DMG-in conformation

Even though no phosphorylation of the CDK8 T-loop is observable its pathway seems positioned within an extensive hydrophilic network involving CycC. The activation of CDK8 by CycC based on a phospho-residue mimicking residue of CycC (E99^{CycC}) has previously been discussed (Hoepfner et al., 2005). In fact within the CDK8 DMG-in conformation of the CDK8/CycC/compound 12 structure, the CycC H3-H4 loop is observed in a relaxed position which increases the CDK8/CycC interaction surface and seems to position E99^{CycC} appropriately. A similar mode of phosphorylation-independent kinase activation has been observed for the CDK6 kinase by complex formation with the virus-encoded cyclin of herpesvirus saimiri (vCyclin). In this complex the open conformation segment of CDK6 is described as very similar to the phosphorylated and fully active CDK2 complex (Schulze-Gahmen et al., 2006).

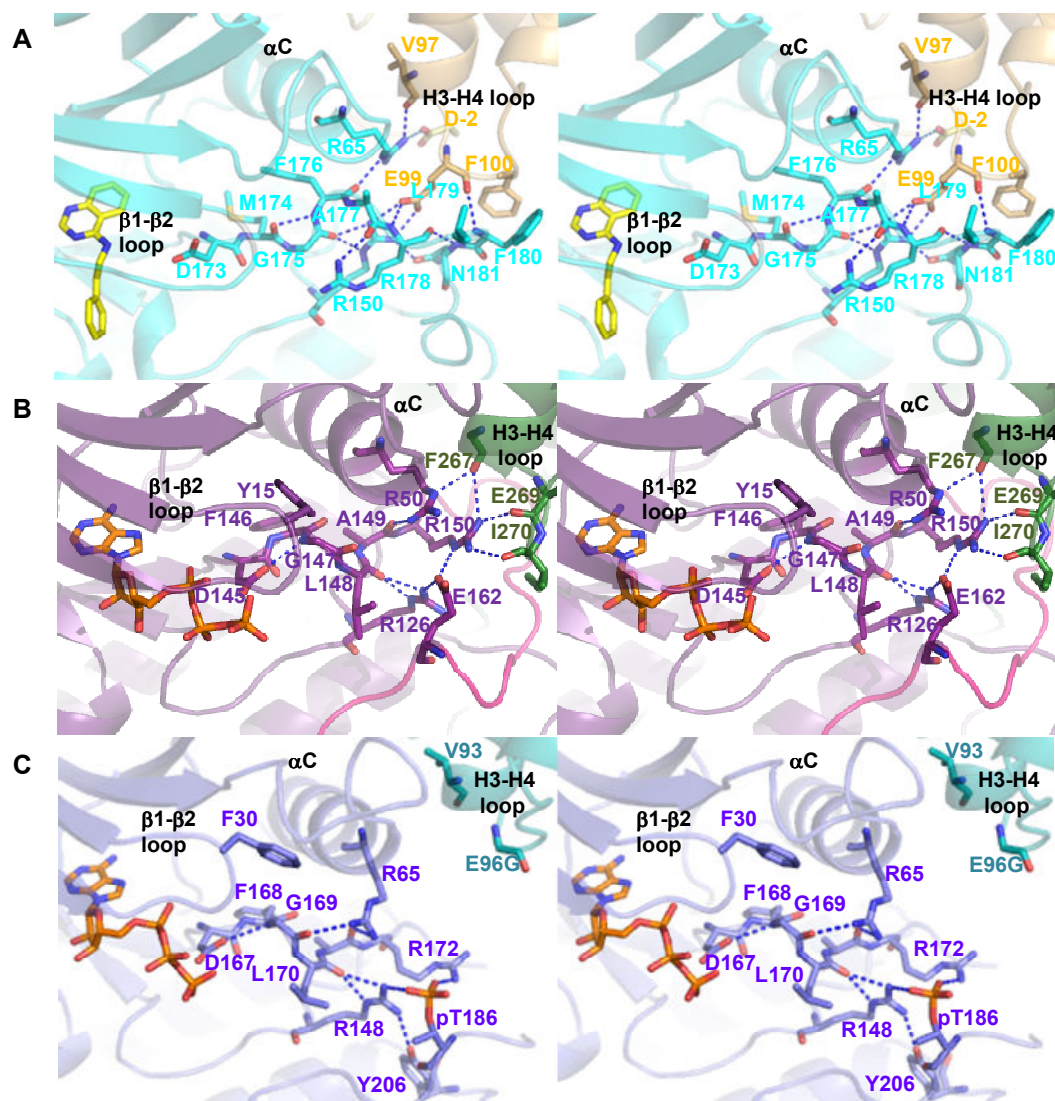


Figure 43 Comparison of the activation segments of CDK8/CycC in the DMG-in conformation to the activation segments of unphosphorylated CDK2/CycA/ATP and phosphorylated CDK9/CycT1/ATP

(A) close-up of the unphosphorylated activation segment of the CDK8 (*turquoise*)/CycC (gold)/ compound 12 (yellow) complex in the DMG-in kinase conformation (4F7S); (B) close-up of the unphosphorylated activation segment of the CDK2 (*plumb*)/ CycA (*forest green*)/ ATP (*deep orange*) complex (1FIN); (C) close-up of the phosphorylated activation segment of the CDK9 (*purple*)/ ATP (*deep orange*) complex (3BLQ);

In comparison with the unphosphorylated CDK2/CycA/ATP complex (1FIN; CDK8/ CDK2 sequence identity: 39.40%, rmsd: 1.72Å) and the phosphorylated CDK9/CycT/ATP complex (3BLQ; CDK8/ CDK9 sequence identity 35.8%, rmsd: 1.63Å) the CDK8 T-loop obviously does not enfold as broadly as observed for the CDK2 and CDK9 T-loops (Figure 43). These divergent pathways of the activation segments are probably caused by the arrangement of the CDK8/CycC complex that influences the position of the CDK8 T-loop which is positioned within a strong interaction network with CycC based

on the CycC H3-H4 loop. Comparing the arrangements of these CDK/cyclin complexes, CDK8/CycC is in an intermediate conformation between the binding angle of CycA to CDK2 and CycT to CDK9. Thereby CycC is rotated by about 10° along its H4 helix relative to CycA and approximately 25° relative to CycT (chapter 4.1, figure 22). As expected the position of the CDK8 DMG-motif is observed in a similar orientation as the DFG-motif of CDK2 and CDK9 (Figure 43A-C). Due to the solvent-molecule bridged H-bond with compound 12, D173^{CDK8} points in a different direction as the corresponding D145^{CDK2} and D167^{CDK9} and does not show the expected H-bond between the D172^{CDK8} δ -oxygen and the main-chain amide of G175^{CDK8} (Kornev et al., 2006). However, this might as well be compensated by the H-bond the G175^{CDK8} backbone-carbonyl forms with the main-chain amide of A177^{CDK8}. Interestingly the CDK8 DMG-motif contacting residue F176^{CDK8} is located beyond the CDK8 α C helix (Figure 43A) whereas the corresponding leucines (L148^{CDK2} and L170^{CDK9}) are solvent-exposed in a rather rigid position given by the T-loop and conditioned by its neighboring interacting residues (Figure 43B and C). Regarding the CDK-wide conserved arginine triad involved in opening the activation segment (Jeffrey et al., 1995; Pavletich, 1999), two of the three conserved arginines of CDK8 (R65^{CDK8} and R150^{CDK8}) contact the exposed side-chain of E99^{CycC} and the V97^{CycC} backbone carbonyl by a hydrophilic network (Figure 43A). Interestingly, the artificially residue D-2^{CycC} of the CycC N-terminus (expression-tag) seems to strengthen the interaction of R65^{CDK8} with the cyclin residues even though an orientation towards E66^{CDK8} is expected. This could also be proven in the following apo structure of CDK8/CycC (see next chapter). Just too distant of the CycC binding surface the main-chain carbonyl of the third conserved arginine, R178^{CDK8} interacts with the main-chain carbonyl of N181^{CDK8}. This interaction obviously serves as another anchor to position the CDK8 T-loop besides the hydrophilic interaction between the E99^{CycC} main-chain carbonyl and the F180^{CDK8} main-chain amide. When comparing with CDK2, E99^{CycC} is found located near E162^{CDK2}, which was reported to compensate for the lack of the phosphate group at the CDK2 T-loop: consequently the CDK2 T loop was found within an at least partially open conformation of the activation segment that allows a correct positioning of ATP even though the CDK2 kinase activity is limited to 1% (Jeffrey et al., 1995, Noble et al., 2005). Within the phosphorylated activation segment of CDK9/CycT, two of the conserved arginines in CDK9 (figure Figure 43C, R172^{CDK9} and R150^{CDK9}) are found within

hydrophilic interaction with the phospho-residue (pT186^{CDK9}) to open the CDK9 activation segment while the third arginine (R65^{CDK9}) is just too distant to form an H-bond. Based on the comparison with CDK2 and CDK9 it is tempting to speculate about an at least partially open conformation of the ordered segment of the CDK8 T-loop which appears induced by the extensive interaction with CycC. Interestingly CycA forms CDK8/CycC similar contacts with CDK2 while CycT is too distant from the binding surface with CDK9. However, this lack of interaction might as well be caused by the mutation of the E99^{CycC} corresponding E96^{CycT} to G96^{CycT} within the respective crystal structure (Baumli et al., 2008). Strengthening the hydrophilic CycC/CDK8 T-loop interaction based on the exposed E99^{CycC} the nearby located F100^{CycC} forms hydrophobic contacts involving F180^{CDK8} and L179^{CDK8} and additionally anchors the CDK8 T-loop conformation. Finally, the still disordered T-loop segment (CDK8 residues 186-196) might be stabilized by binding of an additional activator. Exemplarily MED12 was reported to be necessary for the activation of CDK8/CycC/MED12 as a histone kinase (Knuesel et al., 2009b). Moreover, an activation of CDK8 by binding of a substrate similar to CDK4/CycD (Echalier et al., 2010) might be plausible as the enfolded CycC H3-H4 loop might as well serve as a recruitment patch: it is not only located nearby the CDK8 activation segment but also in neighborhood of the CycC specific groove that was postulated to interact with substrates (Hoepfner et al., 2005; chapter 4.2). Such a substrate-dependent activation of CDK8/CycC could possibly account for the regulation of its activity because in contrast to the timepoint-specific degradation of the cell cycle cyclins, the level of the transcriptional cyclins does not fluctuate significantly throughout the cell cycle (Lolli, 2006). However, activity assays in presence of physiological substrates and/ or further components of the Mediator complex and in parallel to further structural studies will be necessary to find further clues on the catalytically competent activation of the CDK8/CycC kinase.

6.4 The active site of CDK8/CycC in absence of an inhibitor

Up to now it was not clear for the time being whether the CycC-complexed CDK8 kinase preferentially adopts the DMG-in or the DMG-out kinase conformation since only compound-complexed structures of the CDK8/CycC complex were available. Fortunately it was possible to

generate a compound-free CDK8/CycC apo structure (Table 11) by removing compound 12 in several washing steps from the respective crystals.

Data Collection	
Beamline	SLS X06SA
Space group	P 2 ₁ 2 ₁ 2 ₁
	a= 70.18Å; b= 70.94Å; c= 170.42Å
Unit cell dimensions	$\alpha= 90^\circ$, $\beta= 90^\circ$, $\gamma= 90^\circ$
Data Processing	
Resolution (Å)	85.21-2.70
R _{sym}	12.1 [67.9] ^a
Total number of observations	195726 [33346]
Total number unique observations	24071 [4009]
Mean I/sd(I)	15.62 [3.54]
Completeness	99.9 [99.7]
Multiplicity	8.1 [8.3]
Refinement	
Total number residues ^b	536
Number of water molecules	39
R _{Cryst}	21.0
R _{Free} ^c	26.6
r.m.s. bonds (Å)	0.008
r.m.s. angles (deg.)	1.084
Rachmandran plot	
Most favoured geometry (%)	92.4
Additionally allowed (%)	7.6

Table 11 Crystallographic parameters for human CDK8/CycC/apo (to be submitted to the PDB)

^a Values in brackets refer to the highest resolution shell; ^b The final model comprises residues the CDK8 residues 1-359 and CycC residues 1- 264 with three additional residues at the N-terminus of CycC from expression tag (referred to as D-2 D-1 and A0). Segments that are not defined in electron density include the segments 115-122, 185-196, 240-245 and 360-403 for CDK8 and the segment 265-283 for CycC. ^c Test-set contains 3.6 % of measured reflections.

Within this CDK8/CycC/apo structure only buffer-ingredients and solvent-molecules were detectable nearby the CDK8 active site (Figure 44).

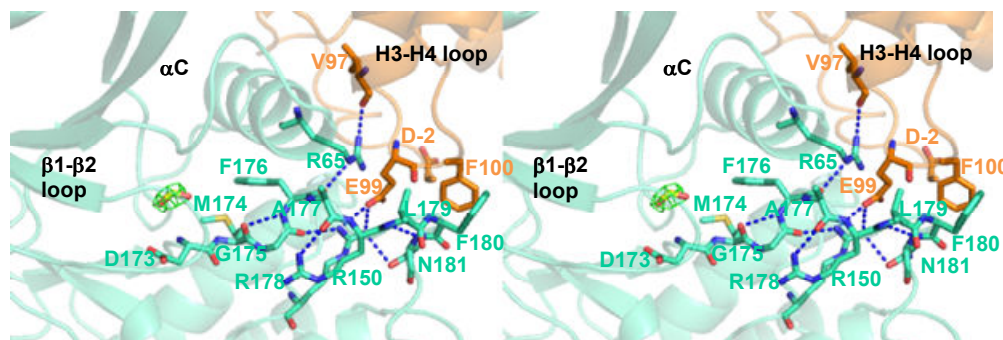


Figure 44 Close up of the activation segment for the CDK8/CycC/apo structure

Close-up of the unphosphorylated activation segment of CDK8 (*teal*) CycC (*orange*) apo complex (to be submitted to PDB); buffer-ingredients yellow; interacting residues and ligands are coloured according to atom type: oxygen red, nitrogen blue, phosphate orange; hydrogen-bonds are indicated with blue lines. F_0-F_C omit map is shown for a buffer-ingredient (formiate) complexed to CDK8/CycC active site with contour level of 3σ (*green*).

In absence of any inhibitor, the CycC-complexed CDK8 kinase is observed within the DMG-in kinase conformation while the position and orientation of the DMG-motif, the conserved arginine triad and other active-site located residues form nearly the same interactions as their corresponding segments and residues within the CDK8/CycC/compound 12 structure (Figure 43A). Interestingly, even in absence of a compound the D173^{CDK8} side-chain adopts the same orientation as in presence of the compound and cannot establish an interaction with the DMG-motif via hydrogen-bonding involving its δ -oxygen. Moreover the CDK8/CycC apo structure reveals that the artificial residue (D-2^{CycC}, expression tag) is rotated outside a binding surface with R65^{CDK8} that is within the same conformation of R65^{CDK8} as observed for the CDK8/CycC/compound 12 structure. Even though the crystal geometry and the crystallization conditions are generally accepted to exert a profound influence on protein conformation within a crystal structure, the here presented CDK8/CycC/apo structure suggests that the CycC-complexed CDK8 kinase preferentially adopts the DMG-in conformation since it has been shown previously that the flexibility within the CDK8/CycC active site allows the induction of a DMG-out kinase conformation by soaking of type II binding compound even within the crystal lattice of CDK8/CycC/type I compound crystals (chapter 5).

In summary, the CDK8 DMG-in kinase conformation within the structures of CDK8/CycC/compound 12 and CDK8/CycC/apo gives hints on a phosphorylation-independent activation and why it is the to-date only known member of the CDK family that enables access to the deep pocket. The available literature indicates an obvious lack of an activatory T-loop phosphorylation which is confirmed in the here presented structures. This seems to be compensated by the extensive interaction of the CDK8 T loop with CycC based on the exposed position of E99^{CycC} as an anchor point. Moreover a pliable conformation of the CDK8 T loop allows the rotation of the DMG-motif embracing residue F176^{CDK8} in addition to D172^{CDK8} and M174^{CDK8} (DMG-motif). Thus the conformation of F176^{CDK8} depicts a secondary characteristic besides the conformation of the DMG-motif being either solvent-exposed (CDK8 DMG-out kinase conformation) or rotated within the deep pocket beyond the CDK8 α C helix (CDK8 DMG-in kinase conformation). Consequently the flexibility of the F176^{CDK8} side-chain possibly enables the CDK8 T-loop to position itself appropriately in order to enable binding capacity within the CDK8 deep pocket accordingly to shape and charge of the respective type II deep pocket binding compound. In contrast the corresponding leucine within all other CDKs (exemplarily L148^{CDK2} and L170^{CDK9}) might not allow a T-loop position that forms a binding surface for the deep pocket binding of a type II compound, especially in presence of an additionally strong fixation as given by the general activatory T-loop phosphorylation. No structure of a type II inhibitor complexed to a T-loop is available yet thus leaving the exact binding mechanism unclear (Liu et al., 2006).

6.5 The crucial influence of CycC binding on the CDK8 conformation allows for selective inhibition of CDK8/CycC and discrimination against the CDK8 conformation in absence of the cyclin

Subsequently a close-up of an overlay of the CDK8/CycC/compound 12 structure (with a CDK8 DMG-in kinase conformation, referred to as 4F7S) to the CDK8/CycC/sorafenib structure (with a CDK8 DMG-out kinase conformation, referred to as 3RGF) elucidates the differences of between type I (compound 12) and type II (sorafenib) binding within the CDK8 active site (Figure 45). In addition to targeting the nucleotide-binding site like compound 12, the type II binding mode of sorafenib expands to the CDK8 deep pocket and causes a DMG-out conformation of the activation segment.

Within the sorafenib binding mode to the CDK8 DMG-out conformation the position of the 3-trifluoromethyl-4-chlorophenyl ring of sorafenib seems to substitute the hydrophobic interactions within the deep pocket that is based on the side-chains of M174^{4F7S} and F176^{4F7S}.

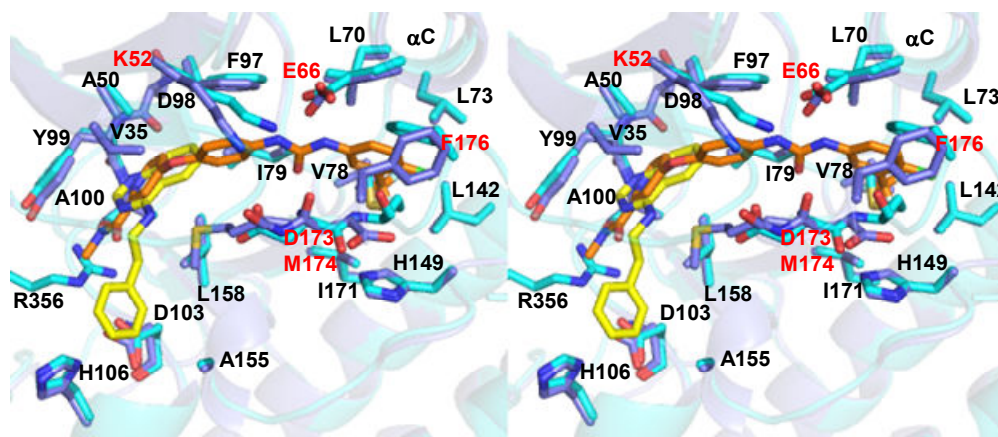


Figure 45 Close-Up of the type I binding mode of compound 12 to the type II binding mode of sorafenib

4F7S: CDK8^{4F7S} turquoise; compound 12 yellow; 3RGF: CDK8^{3RGF} slate blue; sorafenib orange; residues and compounds are colored according to atom type: oxygen red; nitrogen blue; sulfur safron, chloride green, fluorine grey; superposition was centered on the CDKs within COOT, core rmsd 0.56Å; the cyclins are omitted for clarity.

The flip of the DMG motif to the ‘DMG-out’ conformation by rotation of the D173^{3RGF} side-chain induces a translocation of the C α position of M174^{3RGF} by 6.5Å leading to a different pathway of the inactive CDK8 activation segment. While F176^{3RGF} forms extensive hydrophobic contacts with the 3-trifluoromethyl-4-chlorophenyl ring of sorafenib, a large segment of the CDK8 activation segment is displaced and unordered, starting with the side-chain of R178^{3RGF}. Moreover the ion-bridge involving E66^{CDK8} and K52^{CDK8} is displaced by an extensive hydrophobic interaction with sorafenib.

To elucidate the effects of CycC on inhibitor association both compound 12 and sorafenib were screened with the Proteros Reporter Displacement Assay (material and methods, chapter 2.15) against the isolated and the CycC-complexed CDK8. In these assays the ATP-binding site targeting type I inhibitor, compound 12 (Figure 42 and Figure 43A), turned out similarly affine against CDK8 in the presence and absence of CycC (Figure 46: isolated CDK8 $K_d = 2.1\mu\text{M}$; CDK8/CycC $K_d = 2.4\mu\text{M}$). However there is a dramatic difference in the case of sorafenib, whose type II binding mode targets

the conserved E66^{CDK8} of the α C helix and the kinase deep pocket (Figure 45) besides the nucleotide binding site. When screened against isolated CDK8 the affinity of sorafenib is diminished significantly by factor 47 (Figure 46; isolated CDK8: $K_d = 1.88\mu\text{M}$; CDK8/CycC: $K_d = 0.04\mu\text{M}$). On top of that the slow binding kinetics of sorafenib for CDK8/CycC with the extended residence time of 315min is changed to fast binding kinetics with a residence time below the detection level of 1.4min for isolated CDK8.

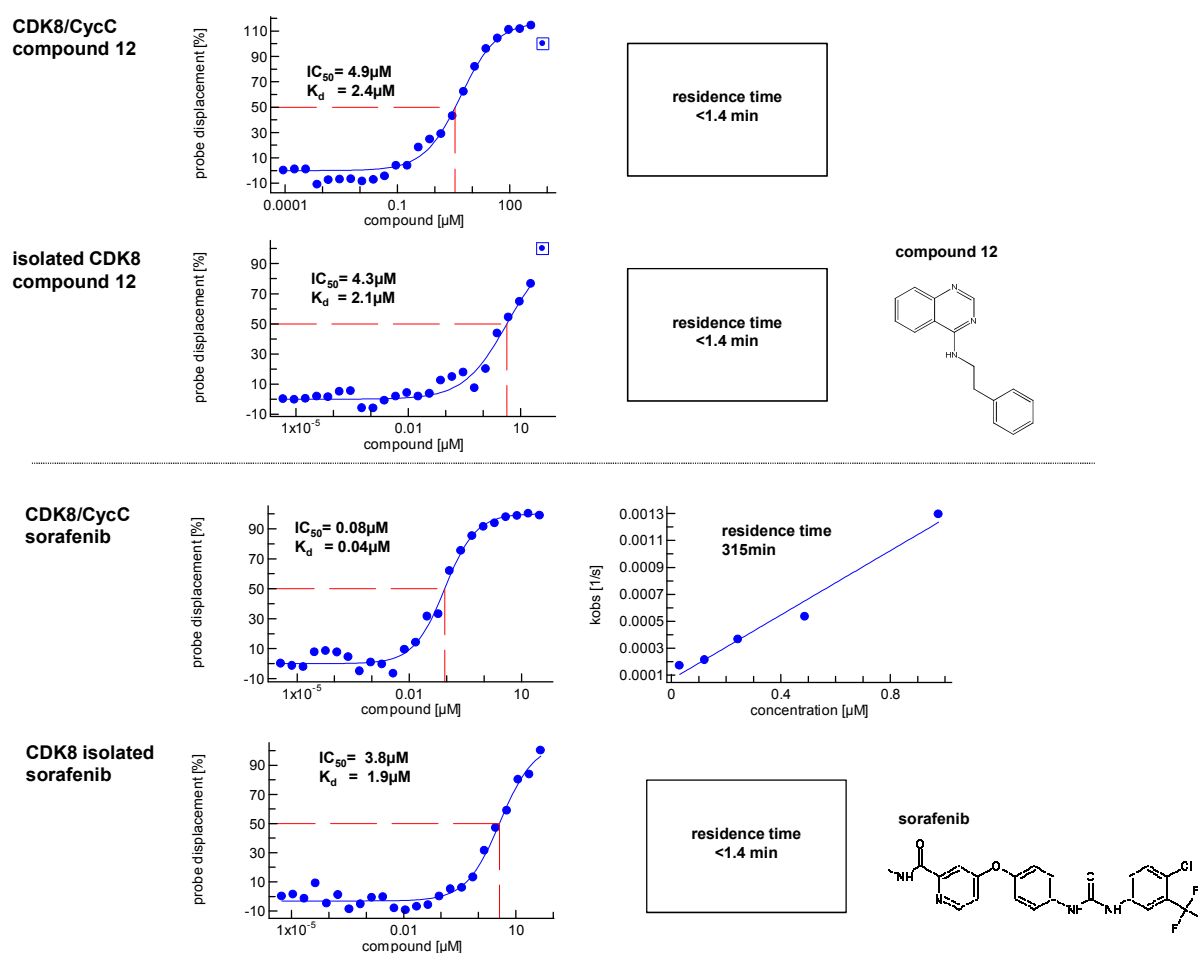


Figure 46 IC_{50} -quantification and determination of binding kinetics of compound 12 and sorafenib for CDK8 in presence and absence of CycC

The IC_{50} measurement of compounds was quantified by regular IC_{50} fitting of the percent reporter displacement values (with no displacement of the reporter corresponding to 0% (C+) and complete displacement (C-) corresponding to 100% after 80min). With the previous determined K_d value of the reporter and its known concentration, the K_d value of the compound is calculated within the equilibrated system from the measured IC_{50} -value by the Cheng-Prusoff equation (material and methods chapter 2.15; equation 2).

These observations are explainable from the crystal structures. As compound 12 performs an ATP-competitive type I binding mode it targets the hinge-region of CDK8 beyond the β 1- β 2 loop (Figure 42 and Figure 43A), that is not within the binding surface of CycC. Therefore the affinity of compound 12 remains constant for both CDK8 conformations as its binding site (comprising the CDK8 hinge-region, the front pocket and the β 1- β 2 loop) does not depend upon binding of the cyclin as also obvious from its crystal structure (4F7S). As sorafenib targets the conserved glutamate of the α C helix within its binding mode (Figure 45) the dramatic decrease of its affinity for isolated CDK8 seems explainable from the general CDK activation mechanism. Thus binding of the cyclin to the α C helix translocates its position towards the CDK activation segment (Jeffrey et al., 1995; Pavletich, 1999). As the urea core of slow binding compounds such as sorafenib forms conserved H-bonds with the conserved glutamate of the α C helix (Dietrich et al., 2010; Namboodiri et al., 2010; chapter 5) this interaction should be at least unfavorable for isolated CDK8 as the α C helix is supposed to be rotated outside the CDK8 activation segment in absence of the cyclin.

6.6 Comparison of the CycC dependence of further SKR compounds

Based on the results of sorafenib and compound 12, parts of the CDK8/CycC specific SKR-compound series (Figure 47 and chapter 5) were screened against isolated CDK8. Compounds 3 - 11 of the SKR-study of CDK8/CycC (chapter 5) were chosen for the following reasons. The unselective compound 6 was chosen as starting point besides to the minimal compound 7 that was shown to bind to the conserved glutamate of the α C helix and the CDK8/CycC deep pocket. Subsequently the “hinge-directed” moiety (nomenclature according to the observed binding mode to CDK8/CycC, chapter 5) of these compound was elongated whereby compound 8, compound 3, compound 9 and compound 4 did not extend far enough to contact the hinge-region of CycC-complexed CDK8 associated with fast binding kinetics (residence time below detection limit < 1.4min). Compound 10 and compound 5 establish one hydrogen-bond with the hinge-region of CycC-complexed CDK8 which is increased to two hydrogen-bonds for compound 11 whereby their binding mode is reconciled within their emerging slow binding kinetics (residence time compound 10: 7min; compound 5: 14min and compound 11: 57min).

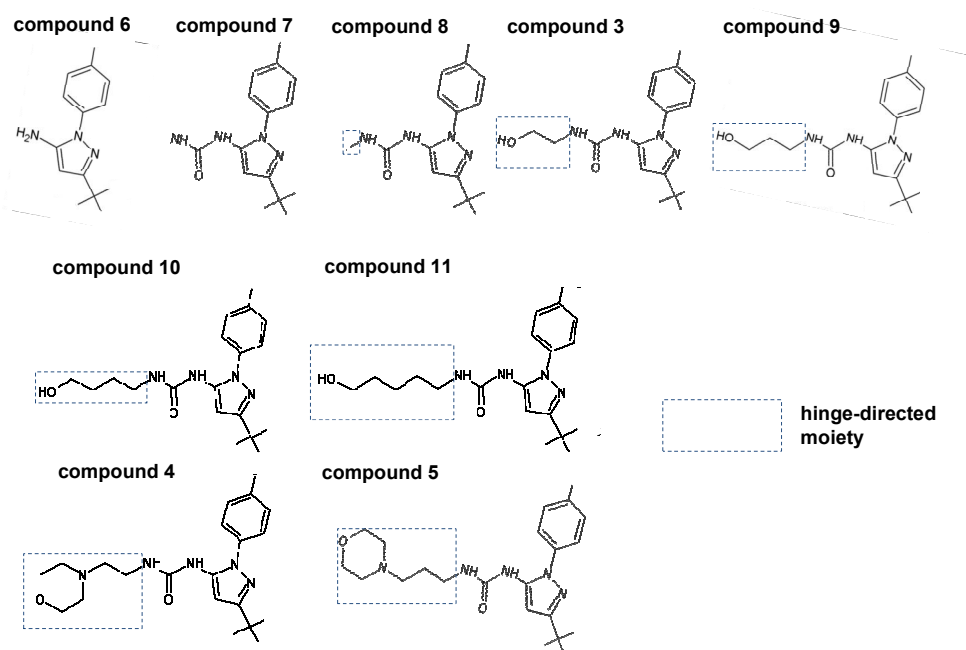


Figure 47 Selected compounds of the CDK8/CycC SKR series

compound 3: 1-[3-tert-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(2-hydroxyethyl)urea; **compound 4:** 1-[3-tert-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-[2-(morpholine-4-yl)ethyl]-urea; **compound 5:** 1-[3-tert-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-[3-(morpholin-4-yl)propyl]urea; **compound 6:** 1-[3-tert-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]amine; **compound 7:** 1-[3-tert-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]urea; **compound 8:** 1-[3-tert-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-methylurea; **compound 9:** 1-[3-tert-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(3-hydroxypropyl)urea; **compound 10:** 1-[3-tert-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(4-hydroxybutyl)urea; **compound 11:** 1-[3-tert-butyl-1-(4-methylphenyl)-1H-pyrazol-5-yl]-3-(5-hydroxypentyl)urea; „hinge-directed“ moiety is boxed blue.

All compounds were measured in triplicates for 80min with the Reporter Displacement Assay (material and methods).

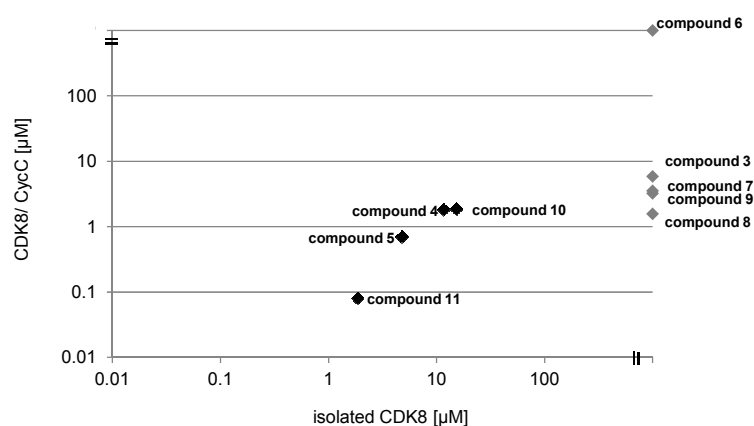


Figure 48 Affinity of the CDK8/CycC SKR compounds for CDK8 in absence and presence of CycC

The K_d -value for isolated CDK8 is plotted on the x-axis, the K_d -value for the CDK8/CycC complex on the y-axis. For the sake of presentability the K_d -value of not applicable compounds was set to $1000\mu\text{M}$.

Measureable affinities were determined only for compound 10 (isolated CDK8: $K_d=15.2\mu\text{M}$; CDK8/CycC $K_d= 1.86\mu\text{M}$), compound 4 (isolated CDK8 $K_d=11.6\mu\text{M}$; CDK8/CycC $K_d= 1.82\mu\text{M}$), compound 5 (isolated CDK8 $K_d= 4.8\mu\text{M}$; CDK8/CycC $K_d= 0.7\mu\text{M}$) and compound 11 (isolated CDK8 $K_d= 1.8\mu\text{M}$; CDK8/CycC $K_d= 0.08\mu\text{M}$) that are significantly diminished in comparison to CDK8/CycC in a range between factor 8 to 22.5 (Figure 48). Obviously, the affinity for isolated CDK8 is only detectable for compounds that include a certain length of their “hinge-directed moiety” as shown for compound 10 (4-hydroxybutyl-urea) and compound 4 (2-(morpholine-4-yl)ethyl]-urea). The affinity for isolated CDK8 increases with the increasing length of the hinge-directed moiety by factor 3 for compound 5 ($K_d= 4.78\mu\text{M}$; 4-hydroxybutyl-urea) and by one order of magnitude for compound 11 ($K_d= 1.86\mu\text{M}$, 5-hydroxypentyl-urea). The diminished affinity of CDK8/CycC specific SKR compounds for isolated CDK8 is also reflected within their binding kinetics. All compounds that show emerging slow binding kinetics for CDK8/CycC have fast binding kinetics for isolated CDK8 (residence time in compound 10: isolated CDK8 below detection limit <1.4min and CDK8/CycC 7min; compound 5: isolated CDK8 below detection limit <1.4min and CDK8/CycC 14min; compound 11: isolated CDK8 below detection limit <1.4min and CDK8/CycC 57min).

In summary the here presented findings underline the possibility to even discriminate between CDK8 in presence and absence of CycC which might be valuable for drug development on this major player in cancerogenesis. However structural studies on isolated CDK8 are necessary to gain further insight on the conformation of CDK8 in absence of CycC before questioning the change of the characteristic slow binding kinetics of sorafenib from slow to fast binding kinetics for isolated CDK8. There are only a few studies available on SKR of type II binders: the exact mechanism of how a compound develops slow binding kinetics remains unclear and is still under discussion. As soon as a structure of isolated CDK8 becomes available it will be of high interest to perform a SKR study similar to the one for CDK8/CycC (chapter 5).

7. References

- Adamczewski, J. P., Rossignol, M., Tassan, J. P., Nigg, E. A., Moncollin, V. & Egly, J. M. (1996). MAT1, cdk7 and cyclin H form a kinase complex which is UV light-sensitive upon association with TFIIH. *Embo J.* **15**, 1877-1884.
- Adler, A. S., McClelland, M. L., Truong, T., Lau, S., Modrusan, Z., Soukup, T. M., Roose-Girma, M., Blackwood, E. M. & Firestein, R. (2012). CDK8 Maintains Tumor Dedifferentiation and Embryonic Stem Cell Pluripotency. *Cancer Res.* **72**, 2129-2139.
- Akoulitchev, S., Chuikov, S. & Reinberg, D. (2000). TFIIH is negatively regulated by cdk8-containing mediator complexes. *Nature* .**407**, 102-106.
- Armache, K. J., Mitterweger, S., Meinhart, A. & Cramer, P. (2005). Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *J Biol Chem.* **280**, 7131-7134.
- Backes, A. C., Zech, B., Felber, B., Klebl, B. & Müller, G. (2008). Small-molecule inhibitors binding to protein kinase. Part II: the novel pharmacophore approach of type II and type III inhibition. *Expert Opin. Drug Discov.* **3**, 1427-1449.
- Barette, C., Jariel-Encontre, I., Piechaczyk, M. & Piette, J. (2001): Human cyclin C protein is stabilized by its associated kinase cdk8, independently of its catalytic activity. *Oncogene.* **20**, 551-562.
- Bártová I, Otyepka M, Kriz Z. & Koca J. (2004). Activation and inhibition of cyclin-dependent kinase-2 by phosphorylation; a molecular dynamics study reveals the functional importance of the glycine-rich loop. *Protein Sci.* **13**, 1449-1457
- Baumli, S., Endicott J. A. & Johnson L. N. (2010). Halogen bonds form the basis for selective P-TEFb inhibition by DRB. *Chem. Biol.* **17**, 931-936.
- Baumli, S., Lolli, G., Lowe, E. D., Troiani, S., Rusconi, L., Bullock, A. N., Debreczeni, J. E., Knapp S. & Johnson, L. N. (2008). The structure of P-TEFb (CDK9/cyclin T1), its complex with flavopiridol and regulation by phosphorylation. *EMBO.* **27**, 1907-1918.
- Betzi, S., Alam, R., Martin, M., Lubbers, D. J., Han H., Jakkaraj, S. R., Georg, G. I. & Schönbrunn, E. (2011). Discovery of a Potential Allosteric Ligand Binding Site in CDK2. *ACS Chem. Biol.* **6**, 492-501.
- Blow, D.M., Wright, C. S., Kukla, D., Rühlmann, A., Steigemann, W. & Huber, R. (1972). A Model for the Association of Bovine Pancreatic Trypsin Inhibitor with Chymotrypsin and Trypsin. *JMB*, **69**, 137-144.

- Bossemeyer, D., Engh, R. A., Kinzel, V., Ponstingl, H. & Huber, R. (1993). Phosphotransferase and substrate binding mechanism of the cAMP-dependent protein kinase catalytic subunit from porcine heart as deduced from the 2.0 Å structure of the complex with Mn²⁺ adenylyl imidodiphosphate and inhibitor peptide PKI (5-24). *EMBO J.* **12**, 849-859.
- Bregman, D. B., Pestell, R. G. & Kidd, V. J. (2000). Cell cycle regulation and RNA polymerase II. *Front Biosci.* **5**, D244-257.
- Casamassimi, A. & Napoli, C. (2007). Mediator complexes and eukaryotic transcription regulation: an overview. *Biochimie.* **89**, 1439-1446.
- Cohen S. X., Morris R. J., Fernandez F. J., Ben Jelloul M., Kakaris M., Parthasarathy, Lamzin V. S., Kleywegt G. J. & Perrakis A. (2004). Towards complete validated models in the next generation of ARP/wARP. *Acta Cryst.* **D60**, 2222-2229.
- Copeland, R. A., Pompliano, D. L. & Meek, T. D. (2006). Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov.* **5**, 730-739.
- Cowtan, K. (1998). Modified phased translation functions and their application to molecular fragment location. *Acta Cryst.* **D54**, 750-756.
- Cowtan, K. (2006). The "Buccaneer" protein model building software. CCP4 newsletter on protein crystallography. Number 44.
- Cramer, P., Bushnell, D. A., & Kornberg, R. D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science.* **292**, 1863-1876.
- Day, P. A., Cleasby, A., Tickle, I. J., O'Reilly, M., Coyle, J. E., Holding, F. P., McMenamin, R. L., Yon, J., Chopra, R., Lengauer, C. & Jhoti, H. (2009). Crystal structure of human CDK4 in complex with a D-type cyclin. *PNAS.* **106**, 4166-4170.
- Devault, A., Martinez, A. M., Fesquet, D., Labbé, J. C., Morin, N., Tassan, J. P., Nigg, E. A., Cavadore, J. C. & Dorée, M. (1995). MAT1 ('menage à trois') a new RING finger protein subunit stabilizing cyclin H-cdk7 complexes in starfish and *Xenopus* CAK. *EMBO.* **14**, 5027-5036.
- Dietrich, J., Hulme, C. & Hurley, L. H. (2010). The design, synthesis, and evaluation of 8 hybrid DFG-out allosteric kinase inhibitors: a structural analysis of the binding interactions of Gleevec, Nexavar, and BIRB-796. *Bioorg. Med. Chem.*, **18**, 5738-5748.
- Dodson, E. J., Winn, M. & Ralph, A. (1997). Collaborative Computational Project, number 4: providing programs for protein crystallography. *Methods Enzymol.* **277**, 620-633.

- Donner, A. J., Szostek, S., Hoover, J. M. & Espinosa, J. M. (2007). CDK8 is a stimulus-specific positive coregulator of p53 target genes. *Mol. Cell.* **27**, 121-133.
- Donner, A. J., Ebmeier, C. C., Taatjes, D. J. & Espinosa, J. M. (2010). CDK8 is a positive regulator of transcriptional elongation within the serum response network. *Nat. Struct. Mol. Biol.*, **17**, 194-201.
- Dynlacht, B. D. (1997). Regulation of transcription by proteins that control the cell cycle. *Nature.* **389**, 149-152.
- Dyson, M. R., Shadbolt, S. P., Vincent, K. J., Perera, R. L. & McCafferty, J. (2004): Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression. *BMC Biotechnol.*, **4**: 32
- Echalier, A., Endicott, J. A. & Noble, M. E. M. (2010). Recent developments in cyclin-dependent kinase biochemical and structural studies. *BBA.* **1804**, 511-519.
- Emsley, P. & Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126-2132.
- Esposito, D., Chatterjee, D. K. (2006): Enhancement of soluble protein expression through the use of fusion tags. *Curr. Opin. Biotechnol.* **17**, 353-358
- Firestein, R., Bass, A. J., Kim, S. Y., Dunn, I. F., Silver, S. J., Guney, I., Freed, E., Ligon, A. H., Vena, N., Ogino, S., Chheda, M. G., Tamayo, P., Finn, S., Shrestha, Y., Boehm, J. S., Jain, S., Bojarski, E., Mermel, C., Barretina, J., Chan, J. A., Baselga, J., Tabernero, J., Root, D. E., Fuchs, C. S., Loda, M., Shivdasani, R. A., Meyerson, M. & Hahn, W. C. (2008). CDK8 is a colorectal cancer oncogene that regulates b-catenin activity. *Nature.* **455**, 547-551.
- Firestein, R. & Hahn, W. C. (2009). Revving the throttle on an oncogene: CDK8 takes the driver seat. *Cancer Res.* **69**, 7899-7901.
- Firestein, R., Shima, K., Noshu, K., Irahara, N., Baba, Y., Bojarski, E., Giovannucci, E. L., Hahn, W. C., Fuchs, C. S. & Ogino, S. (2010). CDK8 expression in 470 colorectal cancers in relation to beta-catenin activation, other molecular alterations and patient survival. *Int J Cancer.* **126**, 2863-73.
- Fisher, R. P., Jin, P., Chamberlin, H. M. & Morgan, D. O. (1995). Alternative mechanisms of CAK assembly require an assembly factor or an Activating Kinase. *Cell.* **83**, 47-57.
- Fujinaga, K., Taube, R., Wimmer, J., Cujec, T. P. & Peterlin, B. M. (1999): Interactions between human cyclin T, Tat, and the transactivation response element (TAR) are disrupted by a cysteine to tyrosine substitution found in mouse cyclin T. *Proc Natl Acad Sci USA.* **96**, 1285-1290

- Garriga, J., Peng, J., Parreno, M., Price, D. H., Henderson, E. E. & Grana, X. (1998). Upregulation of cyclin T1/CDK9 complexes during T cell activation. *Oncogene*. **17**, 3093-3102.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server (Humana Press).
- Gotoh, T., Miyazaki, Y., Sato, W., Kikuchi, K.-I. & Bentley, W. E. (2001): Proteolytic activity and recombinant protein production in virus-infected Sf-9 insect cell cultures supplemented with carboxyl and cysteine protease inhibitors. *Journal of Bioscience and Bioengineering*. **92**, 248-255
- Gouet, P., Courcelle, E., Stuart, D. I. & Metz, F. (1999). ESPript: multiple sequence alignments in PostScript. *Bioinformatics*. **15**, 305-308.
- Haga, A., Hashimoto K., Tanaka N., Nakamura K. T. & Deyashiki Y. (1998): Scalable purification and characterization of the extracellular domain of human autotaxin from prokaryotic cells. *Protein Expr. Purif.*; **59**: 9-17
- Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol*. **11**, 394-403.
- Hannahan, D. & Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*. **100**, 57-70.
- Harper, J. W. & Adams, P. D. (2001). Cyclin-dependent kinases. *Chem. Rev*. **101**, 2511-2526.
- Harper, J. W. & Elledge, S. J. (1998). The role of Cdk7 in CAK function, a retroretrospective. *Genes Dev*. **12**, 285-289.
- He, S.-B., Yuan, Y., Wang, L., Yu, M.-J., Zhu, Y.-B. & Zhu, X.-G. (2011). Effects of cyclin-dependent kinase 8 specific siRNA on the proliferation and apoptosis of colon cancer cells. *J Exp Clin Cancer Res*, **30**:109
- Heitz, F., Morris, M. C., Fesquet D., Cavadore, J. C., Dorée, M. & Divita, G. (1997). Interactions of cyclins with cyclin-dependent kinases: a common interactive mechanism. *Biochemistry*. **36**, 4995-5003.
- Hengartner, C. J., Myer, V. E., Liao, S. M., Wilson, C. J., Koh, S. S. & Young, R. A. (1998). Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Mol Cell*. **2**, 43-53.
- Hoeppner, S., Baumli, S. & Cramer, P. (2005). Structure of the mediator subunit Cyclin C and its implications for CDK8 function. *J. Mol. Biol*. **350**, 833-842.
- Huse, M. & Kuriyan, J. (2002). The conformational plasticity of protein kinases. *Cell*. **109**, 275-282.

- Ikonomou, L., Schneider, Y. J. & Agathos, S. N. (2003). Insect cell culture for industrial production of recombinant proteins. *Appl Microbiol Biotechnol.* **62**, 1-20.
- Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massagué, J. & Pavletich, N. P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature.* **376**, 313-320.
- Jeffrey, P. D., Tong, L. & Pavletich, N. P. (2000). Structural basis of inhibition of CDK-cyclin complexes by INK4 inhibitors. *Genes Dev.* **14**, 3115-3125.
- Kabsch, W. J. XDS. *Appl Crystallogr.* 1993, 795-800.
- Kapoor, A., Goldberg, M. S., Cumberland, L. K., Ratnakumar, K., Segura, M. F., Emanuel, P. O., Menendez, S., Vardabasso, C., Leroy, G., Vidal, C. I., Polsky, D., Osman, I., Garcia, B. A., Hernando, E. & Bernstein, E. (2010). The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. *Nature.* **468**, 1105-1109.
- Kim, S., Xu, X., Hecht, A., & Boyer, T. G. (2006): Mediator Is a Transducer of Wnt/ β -Catenin Signaling. *JBC*; **281**: 14066-14075.
- Knuesel, M. T., Meyer, K. D., Bernecky, C. & Taatjes, D. J. (2009a). The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. *Genes Dev.* **23**, 439-451.
- Knuesel, M. T., Meyer, K. D., Donner, A. J., Espinosa, J. M. & Taatjes, D. J. (2009b). The human CDK8 subcomplex is a histone kinase that requires Med12 for activity and can function independent of mediator. *Mol. Cell. Biol.* **29**, 650-661.
- Kornberg, R. D. (2005). Mediator and the mechanism of transcriptional activation. *Trends Biochem Sci.* **30**, 235-239.
- Kornev, A. P., Haste, N. M., Taylor, S. S. & Eyck, L. F. (2006). Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci U S A.* **103**, 17783-17788
- Krissinel, E. & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2256-2268.
- Kroe, R. R., Regan, J., Proto, A., Peet, G. W., Roy, T., Landro, L. D., Fuschetto, N. G., Pargellis, C. A. & Ingraham, R. H. (2003). Thermal denaturation: a method to rank slow binding, high-affinity P38alpha MAP kinase inhibitors. *J Med Chem.* **46**, 4669-4675.
- Larochelle, S., Pandur, J., Fisher, R. P., Salz, H. K. & Suter, B. (1998). Cdk7 is essential for mitosis and for in vivo Cdk-activating kinase activity. *Genes Dev.* **12**, 370-381.

- Larkin, M. A., Blackshields, G., Brown, N. P., McGettigan P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). ClustalW and ClustalX version 2. *Bioinformatics*. **23**, 2947-2948.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *PROCHECK*: a program to check the stereochemical quality of protein structures. *J App Cryst*. **26**, 283-291.
- Leslie, A. G. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4+ESF-EAMCB Newsletter on Protein Crystallography*, **26**, 48-57.
- Liao, S. M., Zhang, J., Jeffery, D. A., Koleske, A. J., Thompson, C. M., Chao, D. M., Viljoen, M., van Vuuren, H. J. & Young, R. A. (1995). A kinase-cyclin pair in the RNA polymerase II holoenzyme. *Nature* **374**, 193-196.
- Liu, Y. & Gray, N. S. (2006). Rational design of inhibitors that bind to inactive kinase conformations. *Nat Chem Biol*. **2**, 358-364
- Lo Conte, L., Chothia, C. & Janin, J. (1999). The atomic structure of protein-protein recognition sites. *J Mol Biol*. **285**, 2177-2198.
- Lolli, G. & Johnson, L.N. (2007). Recognition of Cdk2 by Cdk7. *Proteins*. **67**, 1048-1059.
- Lolli, G. (2010). Structural dissection of cyclin dependent kinases regulation and protein recognition properties. *Cell Cycle*. **9**, 1551-1561.
- Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). BALBES: a molecular-replacement pipeline. *Acta Crystallogr. D Biol. Crystallogr*. **64**, 125-132.
- Lu, H. & Tonge, P. J. (2010). Drug-target residence time: critical information for lead optimization. *Curr Opin Chem Biol*., **14**, 467-474.
- Malumbres, M. & Barbacid, M. (2005): Mammalian cyclin-dependent kinases. *TRENDS in Biochemical Sciences*, **11**: 630-640
- Marshall, N. F., Peng, J., Xie, Z. & Price, D. H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J Biol Chem*. **271**, 27176-27183
- Martensen, P. M. & Justesen, J. (2001). Specific inhibitors prevent proteolytic degradation of recombinant proteins expressed in High Five cells. *Biotechniques*. **30**, 782-784.
- Meinhart, A., Kamenski, T., Hoepfner, S., Baumli, S. & Cramer, P. (2005). A structural aspect of CTD function. *Genes Dev*. **19**, 1401-1415.

- Morris, E. J., Ji, J. Y., Yang, F., Di Stefano, L., Herr, A., Moon, N. S., Kwon, E. J., Haigis, K. M., Näär, A. M. & Dyson, N. J. (2008). E2F1 represses β -catenin transcription and is antagonized by both pRB and CDK8. *Nature*. **455**, 552-556.
- Murray, A. W. (2004). Recycling the cell cycle: cyclins revisited. *Cell*. **116**, 221-234.
- Müller, G., Sennhenn, P. C., Woodcock, T. & Neumann, L. (2010). The 'retro-design' concept for novel kinase inhibitors. *iDrugs*. **13**, 457-466.
- Namboodiri, H. V., Bukhtiyarova, M., Ramcharan, J., Karpusas, M., Lee, Y. & Springman, E. B. (2010). Analysis of Imatinib and Sorafenib Binding to p38a Compared with c-Abl and b-Raf Provides Structural Insights for Understanding the Selectivity of Inhibitors Targeting the DFG-Out Form of Protein Kinases. *Biochemistry*. **49**, 3611-3618.
- Neumann, L., Ritscher, A., Müller, G. & Hafenbradl, D. (2009). Fragment-based lead generation: identification of seed fragments by a highly efficient fragment screening technology. *J Comput Aided Mol Des*. **23**, 501-511.
- Neumann, L., von König, K. & Ullmann, D. (2011). HTS reporter displacement assay for fragment screening and fragment evolution toward leads with optimized binding kinetics, binding selectivity, and thermodynamic signature. *Methods Enzymol*. **493**, 299-320.
- Nigg, E. A. (1996). Cyclin-dependent kinase 7: at the cross-roads of transcription, DNA repair and cell cycle control? *Curr Opin Cell Biol*. **8**, 312-317.
- Noble, M. E., Endicott, J. A., Brown, N. R. & Johnson, L. N. (1997). The cyclin box fold: protein recognition in cell cycle and transcriptional control. *Trends Biochem. Sci.*, **22**, 482-487.
- Noble, M., Barrett, P., Endicott, J., Johnson, L., McDonnell, J., Robertson, G. & Zawaira, A. (2005). Exploiting structural principles to design cyclin-dependent kinase inhibitors. *Biochim Biophys Acta*. **1754**, 58-64.
- Oelgeschlager, T. (2002). Regulation of RNA polymerase II activity by CTD phosphorylation and cell cycle control. *J Cell Physiol*. **190**, 160-169.
- Otyepka, M., Kríz, Z. & Koca, J. (2002). Dynamics and binding modes of free cdk2 and its two complexes with inhibitors studied by computer simulations. *J Biomol Struct Dyn*. **20**, 141-154.
- Pargellis, C., Tong, L., Churchill, L., Cirillo, P. F., Gilmore, T., Graham, A. G., Grob, P. M., Hickey, E. R., Moss, N., Pav, S. & Regan, J. (2002). Inhibition of MAP kinase by utilizing a novel allosteric binding site. *Nat Struct Biol*. **9**, 268-272.

- Pavletich, N. P. (1999). Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. *J. Mol. Biol.* **287**, 821-828.
- Pflugrath, J. W. (2004): Macromolecular cryocrystallography - methods for cooling and mounting protein crystals at cryogenic temperatures. *Methods.* **34**: 415-423.
- Pinhero, R., Liaw, P., Bertens, K. & Yankulov, K. (2004). Three cyclin-dependent kinases preferentially phosphorylate different parts of the C-terminal domain of the large subunit of RNA polymerase II. *Eur. J. Biochem.* **271**, 1004-1014.
- Rajender, P. S., Vasavi, M. & Vuruputuri, U. (2011). Identification of novel selective antagonists for cyclin C by homology modeling and virtual screening. *Int. J. Biol. Macromol.* **48**, 292-300.
- Regan, J., Capolino, A., Cirillo, P. F., Gilmore, T., Graham, A. G., Hickey, E., Kroe, R. R., Madwed, J., Moriak, M., Nelson, R., Pargellis, C. A., Swinamer, A., Torcellini, C., Tsang, M. & Moss, N. (2003). Structure-activity relationships of the p38alpha MAP kinase inhibitor 1-(5-tert-butyl-2-p-tolyl-2H-pyrazol-3-yl)-3-[4-(2-morpholin-4-yl-ethoxy)naphthalen-1-yl]urea (BIRB 796). *J Med Chem.* **46**, 4676-86.
- Ren, S. & Rollins, J. B. (2004). Cyclin C/Cdk3 promotes Rb-dependent G0 exit. *Cell.* **117**, 239-251.
- Rickert, P., Seghezzi, W., Shanahan, F., Cho, H., & Lees, E. (1996). CyclinC/CDK8 is a novel CTD kinase associated with RNA polymerase II. *Oncogene.* **12**: 2631-2640.
- Rubin, G. M., Yandell, M., D., Wortman J. R., Gabor Miklos, G. L., Nelson, C. R., Hariharan, I. K., Fortini, M. E., Li, P. W., Apweiler, R., Fleischmann, W. Cherry, J. M., Henikoff, S., Skupski, M. P., Misra, S., Ashburner, M., Birney, E., Boguski, M. S., Brody, T., Brokstein, P., Celniker, S. E., Chervitz, S. A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R. F., Gelbart, W. M., George, R. A., Goldstein, L. S. B., Gong, F., Guan, P., Harris, N. L., Hay, B. A., Hoskins, R. A., Li, J., Li, Z., Hynes, R. O., Jones, S. J. M., Kuehl, P. M., Lemaitre, B., Littleton, J. T., Morrison, D. K., Mungall, C., O'Farrell, P. H., Pickeral, O. K., Shue, C., Voshall, L. B., Zhang, J., Zhao, Q., Zheng, X. H., Zhong, F., Zhong, W., Gibbs, R., Venter, J. C., Adams, M. D. & Lewis, S. (2000): Comparative Genomics of the Eukaryotes. *Science*; **287**: 2204-2215.
- Russo, A. A., Tong, L., Lee, J. O., Jeffrey, P. D. & Pavletich, N. P. (1998). Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. *Nature.* **395**, 237-243.
- Rühlmann, A., Kukla, D., Schwager, P., Bartels, K. & Huber, R. (1973). Structure of the Complex formed by Bovine Trypsin and Bovine Trypsin Inhibitor. Crystal Structure Determination and Stereochemistry of the Contact Region. *JMB*, **77**, 417-436

- Sage J. (2004). Cyclin C Makes an Entry into the Cell Cycle. *Dev. Cell.* **6**, 607–608.
- Schmidtke, P., Luque, F. J., Murray, J. B., Barril, X. (2011). Shielded hydrogen bonds as structural determinants of binding kinetics: application in drug design. *J Am Chem Soc.* **133**, 18903–18910.
- Schulman, B. A., Lindstrom, D. L. & Harlow, E. (1998). Substrate recruitment to cyclin-dependent kinase 2 by a multipurpose docking site on cyclin A. *PNAS.* **95**, 10453-10458.
- Schreiber, G. & Keating, A. E. (2011). Protein binding specificity versus promiscuity. *Curr. Opin. Struct. Biol.* **21**, 50-61.
- Schulze-Gahmen, U. & Kim, S. H. (2002). Structural basis for CDK6 activation by a virus-encoded cyclin. *Nat. Struct. Biol.* **9**, 177-181.
- Sherr, C. J. (1996): Cancer Cell Cycles. *Science*; **274**: 1672-1676.
- Simard, J. R., Klüter, S., Grütter, C., Getlik, M., Rabiller, M., Rode, H. B. & Rauh, D. (2009a). A new screening assay for allosteric inhibitors of cSrc. *Nat Chem Biol.* **6**. 394-396.
- Simard, J. R., Getlik, M., Grütter, C., Pawar, V., Wulfert, S., Rabiller, M. & Rauh, D. (2009b). Development of a fluorescent-tagged kinase assay system for the detection and characterization of allosteric kinase inhibitors. *J. Am. Chem. Soc.* **131**, 13286-13296.
- Takaki, T., Echalié, A., Brown, N. R., Hunt, T., Endicott, J. A. & Noble, M. E. (2009). The structure of CDK4/cyclin D3 has implications for models of CDK activation. *Proc Natl Acad Sci U S A.* **106**, 4171-4176.
- Tarricone C., Dhavan R., Peng J., Areces, L. B., Tsai L. H. & Musacchio, A. (2001). Structure and regulation of the CDK5-p25^{nck5a} complex. *Mol Cell.* **8**, 657-669.
- Tassan, J. P., Jaquenoud, M., Leopold, P., Schultz, S. J. & Nigg, E. A. (1995). Identification of human cyclin-dependent kinase 8, a putative protein kinase partner for cyclin C. *Proc Natl Acad Sci U S A.* **92**, 8871-8875.
- Terwilliger, T. C. (2002). Automated main-chain model-building by template-matching and iterative fragment extension. *Acta Cryst.* **D59**, 34-44.
- The PyMOL Molecular Graphics System*, Version 1.2, Schrödinger, LLC, New York, NY.
- Thomas, J. G., Ayling, A., & Baneyx, F. (1997): Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from *E. Coli* to fold or to refold. *Applied Biochemistry and Biotechnology*; **66**, 197-238

- Tummino, P. J. & Copeland, R. A. (2008). Residence time of receptor-ligand complexes and its effect on biological function. *Biochemistry*. **47**, 5481-92.
- Vagin, A. & Teplyakov, A. (2010). Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr*; **66**, 22-25.
- Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F. & Murshudov, G. N. (2004). REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr*. **60**, 2184-2195.
- Viallard, J. F., Lacombe, F., Belloc, F., Pellegrin, J. L., & Reiffers, J. (2001): Molecular mechanisms controlling the cell cycle: fundamental aspects and implications for oncology. *Cancer Radiother*. **2**, 109-29.
- Wallenfang, M. R. & Seydoux, G. (2002). cdk-7 is required for mRNA transcription and cell cycle progression in *Caenorhabditis elegans* embryos. *Proc Natl Acad Sci USA*. **99**, 5527-5532.
- Wan, P. T., Garnett, M. J., Roe, S. M., Lee, S., Niculescu-Duvaz, D., Good, V. M., Jones C. M., Marshall C. J., Springer C. J., Barford, D., Marais, R., & Cancer Genome Project. (2004). Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell*. **116**, 855-867.
- Weinberg, R. A. (1995): The Retinoblastoma Protein and Cell Cycle Control. *Cell*; **81**: 323-330.
- Wilhelm, S., Carter, C., Lynch, M., Lowinger, T., Dumas, J., Smith, R. A., Schwartz, B., Simantov, R. & Kelley S. (2006). Discovery and development of sorafenib: a multikinase inhibitor for treating cancer. *Nat. Rev. Drug Discovery*. **5**, 835-844.
- Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T. & Handa, H. (2006) P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol Cell*. **21**, 227-237
- Yang, Z., Yik, J. H., Chen, R., He, N., Jang, M. K., Ozato, K., Zhou, Q. (2005). Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell*.; **19**, 535-45.
- Zuccotto, F., Ardini, E., Casale, E. & Angiolini, M. (2010). Through the "gatekeeper door": exploiting the active kinase conformation. *J. Med. Chem*. **53**, 2681-2694

Curriculum Vitae

Name: Elisabeth Veronika Schneider
 Geburtsdatum: 10. Juni 1983
 Geburtsort: Altötting/Deutschland
 Nationalität: Deutsch

Akademische Ausbildung

Oktober 2008 bis derzeitig

Doktorarbeit

„Structural and functional studies on human protein kinases exemplified on the human cyclin-dependent kinase 8/ cyclin C complex“

Max Planck Institut für Biochemie
 Proteros Biostructures GmbH
 Prof. Dr. Dr. h.c. Robert Huber

Oktober 2005-Februar 2008

Studium der Ernährungswissenschaft

Technische Universität München

14. Februar 2008

Master Thesis:

Abschluss: Master of Science

“The Impact of Hyperglycemia on NFκB-Signaling in 3T3L1 Cells”

Lehrstuhl für Ernährungsmedizin, TUM

Prof. Dr. Hans Hauner

Oktober 2002-Oktober 2005

Studium der Ernährungswissenschaft

Technische Universität München

02. August 2005

Bachelor Thesis

Abschluss: Bachelor of Science

“Stabilitätsuntersuchungen von Cyanidin und Cyanidin-3,5-Diglucosid”

Lehrstuhl für Biofunktionalität der Lebensmittel, TUM

Prof. Dr. Dr. Gerhard Rechkemmer

September 1993-Juni 2002

König Karlmann Gymnasium Altötting

28. Juni 2002

Hochschulreife

Publikationen

Poster

Strukturelle Vielfalt des Cyanidins und Cyanidin-3,5-diglucosids und die sich daraus ergebenden Konsequenzen für die Analytik;

Brockmeyer, A., Schneider, E., Biemer, E. and Rechkemmer, G.

(43. DGE-Kongress in Stuttgart, Hohenheim, 9.-10. März, 2006)

Residence Time and Thermodynamics. How Novel High Throughput Binding Assays Impact Drug Discovery.

Schneider, E., von König, K., Ullmann, D. and Neumann, L.

(MipTec, Basel, 20.-24.09.2010)

Artikel

Krinninger, P., Brunner, C., Ruiz, P. A., Schneider, E., Marx, N., Foryst-Ludwig, A., Kintscher, U., Haller, D., Laumen, H. and Hauner, H. (2011). Role of the adipocyte-specific NF-κB activity in the regulation of IP-10 and T cell migration. *Am J Physiol Endocrinol Metab.* **300**, E304-311.

Schneider, E. V., Böttcher, J., Blaesse, M., Neumann, L., Huber, R. and Maskos, K. (2011). The structure of CDK8/CycC implicates specificity in the CDK/cyclin family and reveals interaction with a deep pocket binder. *J Mol Biol.* **412**, 251-266.

Schneider, E. V., Böttcher, J., Huber, R., Maskos, K., Neumann, L. (2012). The discovery of slowness in terms of compound binding kinetics: a Structure-Kinetic-Relationship study on CDK8/CycC. *Manuscript submitted.*