Linköping studies in Science and Technology
Dissertation No. 1694

**Structural insights into protein-protein interactions governing
regulation in transcription initiation and ubiquitination**

Madhanagopal Anandapadamanaban

Division of Molecular Biotechnology
Department of Physics, Chemistry and Biology
Linköping University
SE-581 83 Linköping, Sweden

Linköping, 2015

Front cover:

Surface representation of the protein-protein complexes of E2~Ub-TRIM E3 ligase, which I have embellished with interacting proteins/domains in a range of colors.

The author apologizes to those whose work could not be cited directly as original research articles but rather cited the recent reviews, due to space limitations.

All previously published original peer reviewed articles are reprinted with permissions from the publisher.

© Copyright 2015 Madhanagopal Anandapadamanaban, unless otherwise mentioned.

Madhanagopal Anandapadamanaban

Structural insights into protein-protein interactions governing regulation in transcription initiation and ubiquitination

ISBN: 978-91-7685-984-1

ISSN: 0345-7524

Linkoping studies in Science and Technology. Dissertation No. 1694

Printed in Sweden by LiU-Tryck, 2015

*Dream bigger,
Work smarter and harder,
And luck will be your follower*

Principal supervisor:

Prof. Maria Sunnerhagen

Department of Physics, Chemistry
and Biology, Linköping University

Opponent:

Prof. Cynthia Wolberger

Department of Biophysics and
Biophysical Chemistry,
Johns Hopkins University School of
Medicine

Co-supervisors:

Dr. Martin Moche

Department of Medical Biochemistry
and Biophysics, Karolinska Institutet

Examination committee:

Prof. Stig Linder

Department of Medical and Health
Sciences, Linköping University

Prof. Marie Wahren-Herlenius

Department of Medicine,
Center for Molecular Medicine,
Karolinska Institutet

Prof. Maria Selmer

Department of Cell and Molecular
Biology, Uppsala University

Associate Prof. Katja Petzold

Department of Medical Biochemistry
and Biophysics, Karolinska Institutet

Abstract

Virtually every aspect of the cellular processes in eukaryotes requires that the interactions between protein molecules are well coordinated in different regulatory pathways. Any protein dysfunction involved in these regulatory pathways might lead to various pathological conditions. Understanding the structural and functional peculiarities of these proteins molecular machineries will help in formulating structure-based drug design.

The first regulatory process studied here is the RNA polymerase-II mediated transcription of the eukaryotic protein-coding genes to produce mRNAs. This process requires the formation of the ‘transcription initiation’ by the assembly of Pre-Initiation Complex (PIC) formation at a core promoter region. Regulation at this initiation level is a key mechanism for the control of gene expression that governs cellular growth and differentiation. The transcription Factor IID (TFIID) is a conserved multiprotein general transcription factor with an essential role in nucleating the PIC formation, composed of TATA Binding Protein (TBP) and about 14 TBP Associated Factors (TAFs). The here presented crystal structure (1.97 Å) of TBP bound to TAND1 and TAND2 domains from TAF1 reveals a detailed molecular pattern of interactions involving both transcriptionally activating and repressing regions in TBP, thereby uncovering central principles for anchoring of TBP-binding motifs. Together with NMR and cellular analysis, this work provides the structural basis of competitive binding with TFIIA to modulate TBP in promoter recognition.

In eukaryotes, another fundamental mechanism in the regulation of cellular physiology is the post-translational modification of substrate proteins by ubiquitin, termed ‘ubiquitination’. Important actors in this mechanism are the ubiquitin-ligases (E3s) that culminate the transfer of ubiquitin to the substrate and govern the specificity of this system. One E3 ligase in particular, TRIM21, defines a subgroup of the Tripartite Motif (TRIM) family, which belongs to the major RING-type of E3 ubiquitin ligases, and plays an important role in pathogenesis of autoimmunity by mediating ubiquitination of transcription factors. The crystal structure (2.86 Å) of the RING domain from TRIM21 in complex with UBE2E1, an E2 conjugating enzyme, together with the NMR and SAXS analysis as well as biochemical functional analysis, reveals the molecular basis for the dynamic binding interfaces. The TRIM21 mode of ubiquitin recognition and activation for catalytic transfer of ubiquitin can be modeled onto the entire TRIM family.

Finally, we explored the concepts of conformational selection in proteins as a possible key component for protein-mediated transcriptional regulation. In this framework,

MexR, a bacterial repressor of the MexAB-OprM efflux pump, and its mutant Arg21Trp were studied as an example for proteins presenting different conformations. The residue Arg21Trp mutation is clinically identified to cause of Multi-Drug Resistant (MDR) by attenuated DNA binding, and leads to the overexpression of the MexAB-OprM efflux pump. With the crystal structure (2.19 Å) of MexR mutant Arg21Trp, in combination with MD-simulations and SAXS for both wild-type and mutant, we could unravel the atomic details of the wild-type conformations consisting in subsets of populations of DNA bound and unbound forms. Remarkably, the mutant Arg21Trp stabilize the DNA unbound state and shifts MexR in a pre-existing equilibrium, from a repressed to a derepressed state.

Taken together, these studies substantially broaden our knowledge at a molecular level in protein interactions that are involved in transcriptional regulation and ubiquitination, studied by a carefully selected combination of complementary structural methods spanning different resolutions and time scales.

Populärvetenskaplig sammanfattning

Proteiner är fascinerande molekylära beståndsdelar som finns i varje levande cell, där de utför en mängd olika funktioner som är centrala för liv. Alla vitala cellulära processer kräver att proteinernas samverkan är väl koordinerad, så att rätt sak utförs i rätt tid, och på rätt plats. Felaktigheter i dessa huvudvägar för livsfunktionerna kan leda till livshotande tillstånd som cancer och autoimmuna sjukdomar. Därför är det av stort värde att vi lär oss förstå hur proteins struktur och funktion hänger samman, idealt sett på atomär nivå. Denna kunskap är nödvändig för att vi ska kunna skapa nya specifika läkemedel som potentiellt kan styra om interaktionsvägarna och korrigera defekta proteinfunktioner.

I min forskarutbildning har jag studerat tre proteiner med fokus att förstå den molekylära basen för proteininteraktioner och hur detta är kopplat till funktion och reglering: i) det DNA-bindande proteinet TBP och dess interaktioner med TAF1, som modularar det essentiella steget transkriptionsinitiering, ii) proteinet TRIM21, som styr interferonreglering och är ett mål för autoantikroppar hos patienter med Sjögrens syndrom, iii) det bakteriella proteinet MexR, som reglerar uttrycket av de pump-proteiner som gör att patogena organismer kan överleva trots närvaro av antibiotika och där mutationer leder till antibiotikaresistens. För att studera proteinernas mekanismer på olika tidsskalor och vid olika upplösning har jag använt flera metoder: kristallografi, som ger hög upplösning men av statiska strukturer, NMR (kärnmagnetisk resonans), där man kan studera dynamiska proteiner och proteininteraktioner i lösning, och SAXS (låginkelröntgen), som är idealiskt för studier av större proteinkomplex. Tillsammans med molekylodynamiska simuleringar och biokemiska funktionsstudier har jag kunnat studera samverkan mellan proteiner och deras partners på molekylär nivå inom områdena transkriptionsreglering och ubiquitinerings. Resultaten ökar förståelsen för hur den molekylära och dynamiska komplexiteten i dessa interaktioner styr livsfunktionerna.

LIST OF PUBLICATIONS AND MANUSCRIPTS

Paper I.

Anandapadamanaban, M., Andresen, C., Helander, S., Ohyama, Y., Siponen, M., Lundström, P., Kokubo, T., Ikura, M., Moche, M., Sunnerhagen, M. (2013) High-resolution structure of TBP with TAF1 reveals anchoring patterns in transcriptional regulation. *Nature Structural and Molecular Biology* 20(8): 1008-1014

Paper II.

Anandapadamanaban, M.*, Pilstål, R.*, Andresen, C., Trewhella, J., Moche, M., Wallner, B., Sunnerhagen, M.

Population shift disengages DNA binding in a multidrug resistance MexR mutant. (2015, *Manuscript in submission*)

* These authors contributed equally.

Paper III.

Espinosa, A.*, Hennig, J.*, Ambrosi, A., **Anandapadamanaban, M.**, Abenius, M.S., Sheng, Y., Nyberg, F., Arrowsmith, C.H., Sunnerhagen, M., Wahren-Herlenius, M., (2011) Anti-Ro52 autoantibodies from patients with Sjögren's syndrome inhibit the Ro52 E3 ligase activity by blocking the E3/E2 interface. *J Biol Chem.* 286(42): 36478-91

Paper IV.

Anandapadamanaban, M., Kyriakidis, N., Espinosa, A., Round R.A., Trewhella, J., Wahren-Herlenius, M., Moche, M., Sunnerhagen, M.

Structure of a TRIM21-UBE2E1 complex reveals the specificity of E2 and ubiquitin recognition by TRIM E3 RINGS

(2015, *in Manuscript*)

Additional peer-reviewed, original research publications not included in the thesis

Löw, C., Quistgaard, E., Kovermann, M., **Anandapadamanaban, M.**, Balbach, J., Nordlund, P. (2014). Structural basis for PTPA interaction with the invariant C-terminal tail of PP2A. *Biological Chemistry*, 395(7-8), pp. 685-911

Dovega, R., Tsutakawa, S., Quistgaard, E.M., **Anandapadamanaban, M.**, Löw, C., Nordlund, P. (2014) Structural and biochemical characterization of human PR70 in isolation and in complex with the scaffolding subunit of protein phosphatase 2A. *PLoS ONE* 9(7): e101846

Helander, S.*, Montecchio, M.*, Pilstål, R.*, Su, Y.*, Kuruvilla, J., Elvén, M., Ziauddin, J.M.E., **Anandapadamanaban, M.**, Cristobal, S., Lundström, P., Sears, R., Wallner, B., Sunnerhagen, M. Pre-anchoring of Pin1 to unphosphorylated c-Myc in a fuzzy complex regulates c-Myc activity. *Structure* 2015, in press

CONTRIBUTION REPORT

Madhanagopal Anandapadamanaban (M.A.)

PAPER I.

(1st author) - M.A., prepared the fusion complex of TBP with TAF1 recombinant protein for X-ray crystallography structure determination, and the deuterated ²H, ¹³C, ¹⁵N sample for NMR measurements. M.A., performed the crystallization and diffraction, solved the crystal structure by molecular replacement, and extended a previously published heteronuclear NMR backbone resonance assignment. M.A., carefully evaluated the crystal structure and integrated this with the joint results of all co-authors, and contributed significantly to the writing of all sections in the paper.

PAPER II.

(Shared 1st author) - M.A., cloned the MexR Wild-type and R21W mutant, produced the recombinant protein for X-ray crystallography and performed and evaluated SAXS measurements. M.A., produced highly diffracting crystals of MexR-R21W, determined its crystal structure, performed the NMR condition screening, and was in collaboration highly involved in analyzing the MD simulations data. M.A., contributed extensively to the writing of the manuscript, jointly with the shared 1st author R.P.

PAPER III.

(3rd co-author) - M.A., was involved in the later stage of this article. M.A purified ¹⁵N RING domain from Ro52/TRIM21, repeated the NMR binding experiments of P6 antibody titration to the RING domain, and participated in the evaluation of the results by NMR chemical shift perturbation analysis.

PAPER IV.

(1st author) – Throughout his thesis M.A., was deeply involved in project design and independently contributed with strategies that he then successfully pursued. M.A., cloned, expressed and purified all proteins required for the analysis. He performed the UBE2E1-TRIM21 complex experiments using integrated structural biology techniques of X-ray crystallography, NMR spectroscopy and SAXS, in combination with function related hydrolysis assays. M.A., wrote the entire first version of the manuscript, which was then jointly revised with co-authors.

ABBREVIATIONS

DNA	Deoxyribonucleic acid
aa	Amino acids
RING	Really Interesting New Gene
TBP	TATA-box Binding Protein
TFIID	Transcription Factor II D
MDR	Multi-Drug Resistance
MarR	Multiple antibiotic repressor Regulator
SAL	Salicylate
Ub	Ubiquitin
UBL (Ubl)	Ubiquitin-like modifier
polyUb	Polyubiquitin
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin-ligase enzyme
UBD	Ubiquitin-binding domains
UPS	Ubiquitin-proteasome system
DUBs	Deubiquitinating enzymes
TRIM	TRi partite Motif
RBCC	RING-B-box-Coiled-coil
CC	Coiled-coil
NMR	Nuclear Magnetic Resonance
SEC	Size-Exclusion Chromatography
AGF	Analytical Gel Filtration
IMAC	Immobilized Metal Affinity Chromatography
MD	Molecular Dynamics simulations
CD	Circular Dichroism spectroscopy
NMR	Nuclear Magnetic Resonance
HSQC	Heteronuclear Single Quantum Coherence
TROSY	Transverse Relaxation-Optimized Spectroscopy
SAXS	Small Angle X-ray Scattering

Table of contents

Preface	XIII
1. Introduction to protein structure and regulation	1
1.1. Protein structure and function	1
1.2. Protein-protein interaction	2
2. Introduction to the RNA Pol-II transcription initiation	5
2.1. RNA Pol-II transcription initiation.....	5
2.2. Structural features of TBP.....	7
2.3. Regulation of TBP and its interplay with interactors	7
2.4. TFIID activity modulated by gene-specific transcription activators	13
3. Introduction to the MarR transcriptional regulators	15
3.1. MarR, a regulator of multiple antibiotic resistances	15
3.2. Structural features of MarR homologues	15
3.3. MexR, a MarR repressor of the <i>P. aeruginosa</i> efflux pumps.....	20
4. Introduction to protein ubiquitination	25
4.1. The ubiquitination system and ubiquitin-binding proteins	25
4.2. Structural features of Ub	28
4.3. E2 ubiquitin-conjugating enzyme	30
4.4. RING-containing E3 ligases	33
4.5. E2 - RING E3 interface	34
4.6. Structures of E2~Ub-E3	39
4.7. Current clinical drug targets in ubiquitination system	41
4.8. TRIM family of RING-type E3 ubiquitin-ligases	42
4.9. TRIM21 -the cellular role	43
4.10. TRIM interaction with subsets of E2s	44
5. Methodology	49
5.1. Recombinant protein production in <i>E.coli</i>	49
5.2. Integrated structural techniques	50
Summary of Papers	57
Prospective	61
Acknowledgments	63
References	66

Preface

The central theme in this current thesis is to structurally underpin the mechanistic details of protein-protein interaction and regulation in transcriptional initiation and ubiquitination. Malfunctions in these processes are being associated with a spectrum of human diseases. Therefore, finding the molecular peculiarities that exist with these regulatory macromolecules is essential both for understanding of biology and for rational drug design.

This thesis is the summary of my work as a graduate student and illustrates the findings as **Papers I-IV**. In order to provide basic understanding of the Papers, I provide a brief introduction of the respective fields as illustrated below:

Chapter 1 gives basic introduction and an overview of protein structure and interaction.

Chapter 2 introduces the TBP protein, for the **Paper I** findings.

Chapter 3 introduces the family of MarR transcriptional repressors, in particular MexR, and serves as a basis for **Paper II** findings.

Chapter 4 describes the ubiquitination system and the TRIM RING E3 ligases, in order to address the **Paper III & Paper IV** work.

Chapter 5 briefly describes the combined methodology used in this thesis, to obtain the results of above-mentioned research Papers.

Lastly, the summary of Papers and future outlook are presented.

Chapter 1.

Introduction to protein structure and regulation

Biological systems

Important cellular functions, such as cell growth, division and programmed cell death, primarily rely on proteins, which are macromolecules encoded by the genetic material called DNA. The process providing the framework of the conversion of the genomic information into proteins is divided in two stages and is defined as the ‘central dogma of biology’¹. Transcription is the first step, consisting of the production of the mRNA from a DNA template. Protein translation, which is carried out by the ribosomes in the cell cytoplasm, is the ensuing process of translating mRNA into a protein. Furthermore, the cellular proteins at the end of their lifespan are degraded by the ubiquitin-proteasome system, through the ubiquitination pathway ensuring protein turnover. Proteins directly control the cell through the chemical reactions they perform, by adopting various structural conformations, and by performing specific interactions in the cellular environment to accomplish a task. Therefore, to understand the mechanistic function of cells, it is important to identify the function of its key players: the proteins.

1.1. Protein structure and function

The word ‘protein’ was first mentioned by the Swedish chemist Jöns Jacob Berzelius in one of his letters to the Dutch chemist Gerardus Johannes Mulder, and subsequently it became the term currently in use for these macromolecules with a wide array of functions and tasks in the cell². We now know that proteins play innumerable roles within biological functions, from catalyzing chemical reactions to building structures of all living things. In apparent contrast to this wide range of function, proteins are made up of only 20 different building blocks called the amino acids. The ways these amino acids are arranged determine the final shape and function of the protein. The amino acids are joined by peptide bonds, which link their amino- and carboxyl groups in a linear sequence defined as the primary structure (**Fig. 1**). Protein chains often fold into secondary structures, stabilized by hydrogen bonds: they can form α -helices, when a protein chain twists back onto itself, forming a regular pattern of hydrogen bonds between the backbone atoms of nearby amino acids, or alternatively the backbone atoms of the chains can interact side by side to form β -sheets. Many proteins fold into more complex shapes called the tertiary structure with their hydrophobic side chains sheltered inside away from the surrounding water. Cofactors may help the stabilization of special folds, as for the zinc-finger proteins. For

example, the RING (Really Interesting New Gene) domain is a zinc-finger domain playing a key role in the ubiquitination pathway in the protein E3 ligase. The RING domain comprises 40-60 amino acids as primary sequence, and 16% of α -helix and 24% of β -sheet, while the remaining are in loop regions (**Fig. 1**). Furthermore, the RING includes eight specific cysteines (Cys) and histidines (His) residues, that together coordinate two zinc ions. Consequently, in this RING domain structure the α -helix and β -sheets are brought together forming the tertiary structure by the two zinc ions (**Fig. 1**). The formation of quaternary structures employs functional regulation-mediated associations, such as homo or hetero-dimer to oligomer complexes, thereby facilitating the protein to achieve its function.

Figure 1: Hierarchical nature of protein structure formation. Primary protein sequence composed of amino acids (shown as beads) that are connected via a peptide bond. Three-dimension (3D) -Tertiary structure is formed by the secondary structural elements like α -helix (orange) and β -sheets (magenta) (inset) that fold into a compact globule structure. The example presented here is the RING domain from TRIM21 that binds two Zinc (Zn^{2+} , shown as black spheres) ions to forms its tertiary structure (PDB code- 5A4S, **Paper IV**). Lastly, the quaternary structure formations are based on the subunit interactions that are functionally relevant; here-presented RING domain forms homodimer.

1.2. Protein-protein interaction

The biological effects of a protein molecule depend not only on the protein itself, but also on its physical interaction with other molecules. The interacting molecules that can be recognized include DNA, RNA and small molecule ligands, but also other proteins, which is referred to as Protein-Protein Interaction (PPI). The PPI interfaces between proteins can either form through complementary surface molecules of two proteins, or by the flexible loop region from one protein contacting the structurally well-defined surface of another protein. The latter mode of interaction enables the flexibility within the recognition molecule to be modulated upon binding, possibly

from being a disordered region leading to the formation of a secondary structure, thereby increasing the selectivity. For example, the disordered domain TAND1 from TAF1 forms an α -helical secondary structure element upon interaction with the TBP protein, while specific disordered recognition regions of TAND2 interact with TBP as well (**Paper I**). This observed disorder-to-order structural change in the TAND1 region of TAF1 could be explained as a phenomenon of ‘folding upon binding’ to their partner, and makes it possible to obtain high specificity also with low binding affinities³. Some protein molecules are also regulated by chemical modifications referred as post-translational modifications (PTMs), which increase the diversity of the proteome. Among these, the most studied modifications are protein phosphorylation and ubiquitination. Protein phosphorylation is the attachment of a phosphoryl group at specific sites; this could cause a conformational change or create a negative charged group that can be repulsive or attractive for an interaction and thereby modulate the PPI surface. In the context of ubiquitination, the attachment of ubiquitin onto the desired substrate protein at specific sites as well as the type of ubiquitin processivity dictates the fate of the ubiquitinated substrate proteins. Moreover, some proteins undergo cross-talk between different modifications, as well as reversible modifications, thus indicating the presence of additional mechanisms that facilitate specific regulation.

Concept of conformational selection and population shift in protein interactions

In early days, the molecular recognition of protein-protein, protein-DNA and/or protein-small molecule interactions were viewed as rigid bodies, with a ‘lock-and-key’ analogy that required a structurally complementary fit of the rigid molecules⁴, similar to the fit of puzzle pieces. However, with increased understanding that protein molecules are flexible and dynamically altered, the ‘induced-fit’ model was proposed⁵. In this model, the interacting ligand or protein partner induces conformational changes in the target protein, and the process of recognition is achieved primarily in two states, defined as ‘open’ and ‘closed’ states. The interacting partner binds to the open conformation and as a consequence of conformational changes, results in a closed state conformation. But with an in-depth analysis of the dynamic fluctuation associated within the proteins themselves and upon interactions, it is evident that there is much more to be defined than the strictly posing ‘open’ and ‘closed’ conformation states or as a binary switches of proteins for molecular recognition. In other words, there are more conformational states that are intermediately accessed in macromolecular complex systems. In fact, it is becoming more widely recognized that proteins in their absence of effector or ligand or interacting protein (referred as ‘apo’) often exist as an ensemble of all the possible

conformations, and upon interacting with another molecule with their conformations that are compatible to bind (‘conformational selection’), the pre-existing conformational states are redistributed and the equilibrium is shifted towards the favorable selected conformations (‘population shift’)⁶. However, the optimizations of interactions at the molecular level are complemented through induced fit within these complexes^{6,7}.

The best example for the above mentioned concept from this study includes MexR, which is a DNA repressor protein. MexR includes the pre-existing conformational states of DNA-compatible, -incompatible and intermediate structural conformations in its apo form (**Fig. 2**). The preferential interaction of MexR’s higher-energy conformations to a small molecule, such as a ligand, or selective mutations on itself, causes a shift in the distribution of its accessible conformational states (**Paper II**). In another example of conformational selection, ubiquitin thioester linked with E2 molecule (E2~Ub) adopts a range of E2~Ub conformations⁸. Interestingly, upon interaction with the RING domain, that mediates the ubiquitin recognition, the equilibrium is shifted towards favorable E2~Ub conformations and thereby the process progresses to the next step of catalytic ubiquitin transfer (**Paper IV**).

Figure 2. Schematic illustrations of conformational selection and population shift (Paper II).

Taken together, the molecular recognition and binding interfaces presented in large cellular network regulate different functional pathways that are associated for certain mechanism. How these processes are dysregulated during diseases has been the research focuses for the last three decades in structural biology. Therefore, molecular insights of the interfaces, together with the dynamics associated with the complex systems, would provide comprehension of the structure and function of these regulatory components and possible drug targets for specific diseases.

Chapter 2.

Introduction to the RNA Pol II-transcription initiation complex

-With a focus on structural insights into the TBP subunit of TFIID

Transcription of the eukaryotic genome is a highly coordinated and an intricately regulated process, carried out by three functionally distinct classes of nuclear DNA-dependent RNA Polymerases (Pols): RNA Pol I, II and III. While RNA Pol I and Pol III transcribe genes encoded for ribosomal RNA and small non-coding RNA (tRNA and 5sRNA), respectively, RNA Pol II transcribes protein-coding genes to mRNAs (reviewed in ref ⁹). Transcription proceeds through three distinct stages: initiation, elongation and termination. Among these, the initiation stage of gene transcription by RNA Pol II is a highly regulated process for the protein-coding gene expression, especially important for cell growth and differentiation. The step-wise pathway of transcription initiation includes: (i) the binding of gene-specific regulatory factors near the initiation site, which recruits the transcription machinery, (ii) a pre-initiation complex (PIC) formation by the assembly of general transcription factors and RNA Pol II at the promoter site, (iii) unstable open complex formation by DNA strand unwinding and (iv) RNA synthesis by Pol II and final release from the promoter region¹⁰. In the entirety of this process, the most critical step is the formation of macromolecular PIC complex and the assembly that governs the core promoter recognition complex to specific target genes. The last two decades have provided extensive structural insights into transcription initiation (for recent review ref ¹¹), but we still lack knowledge on the molecular principles governing the assembly of the transcriptional complexes.

2.1. RNA polymerase II transcription initiation – A structural view

In the RNA Pol II mediated transcription initiation, Pol II associates with general Transcription Factors (TF) such as TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIF, at the gene promoter DNA for formation of the PIC macromolecular complex⁹. The nuclear gene promoters such as core promoter, promoter proximal and distal enhancer elements, contain combinations of DNA sequences, which are specifically recognized by a distinct subsets of transcription factors in order to mediate RNA Pol II gene transcription. The transcription factor TFIID plays a central role in the recognition of the core promoter and nucleates the formation of the PIC complex at the

transcriptional start site (TSS) by recruiting other general transcription factors¹¹. The first eukaryotic core promoter to be identified was the TATA promoter element, which is an A/T-rich sequence located approximately 25 to 30 nucleotides upstream of the transcription start site. The TATA-box binding protein (TBP) subunit of TFIID recognizes the consensus TATA box and is essential for nucleating the PIC complex assembly. The flanking regions of TATA-box promoter elements contain TFIIB transcription factor recognition elements such as the BRE promoter element, while other core promoters such as Inr, MTE, DPE and DCE, which are recognized by the specific TBP- associated factors (TAFs) within the TFIID macromolecular complex⁹. The presence of additional core promoters other than the TATA promoter, confer the functions of transcriptional initiation in TATA-less promoters.

In the context of general transcription factors, the sequential processes of transcription initiation involves the assembly of TBP subunit of TFIID at promoter DNA, followed by the TFIIA, TFIIB, Pol II and TFIIF complexes, resulting in the core initiation complex, which recruits the complex formation with TFIIE and TFIIH, for the functional DNA unwinding and thereafter for the elongation complex. The structural architecture of the core initiation complex assembly was recently determined, and more specifically, illustrates how TBP and TFIIB cooperatively function to load the promoter DNA onto Pol II transcription machinery¹².

TFIID assembly and architecture

The initiation of gene transcription on the respective core promoter elements by RNA pol II is highly ordered and requires RNA pol-specific initiation factors, such as TFIID. This auxiliary transcription factor of transcription initiation is a large macromolecular complex of approximately 1.2 MDa in size, is composed by TBP and a group of 13-14 conserved TBP -associated factors (TAFs) (for recent TFIID structural review, see ref ¹³). Structural studies of the TFIID core complex have elucidated the three-dimensional architecture of the complex showing TAF association and the addition of a unique TAF8-TAF10 complex leading to a functional asymmetric feature which nucleates the holo-TFIID, and thereby assembly begins¹⁴. In TFIID, the TATA-box binding protein (TBP) is an essential subunit and a universal component common to all three distinct RNA Pols¹⁵, although distinct functional general transcription factors are required for driving specific RNA Pols machinery. In other words, TBP is an integral subunit for the transcription factors of SL1, TFIID and TFIIB, which are specific initiation core factors to RNA Pol I, II and III, respectively⁹. TBP recruitment to a promoter DNA for gene expression is a key for regulation of transcription activation, and therefore a crucial step in transcription initiation. TBP-binding to the TFIID complex leads to a major structural molecular

assembly rearrangement, contributing to the dynamic association of complexes to modulate the shape, composition and activity for gene transcription¹⁶.

2.2. Structural features of TBP

The core domain of TBP structure adopts a molecular ‘saddle-shape’ pertaining to an intramolecular domain symmetry (pseudo-two-fold). This approximate two-fold symmetry is generated by each five-stranded antiparallel β -sheet and two α -helices (H1 and H2, where H1’ and H2’ refers to the symmetry helices), furthermore, the large loops connecting the initial β -sheets are N and C lobe (N-terminal and C-terminal lobes, also referred as ‘stirrups’)¹⁷ (**Fig. 3a**). The two major TBP surfaces that are fundamental for the complex formation are the concave and convex surface (**Fig. 3a**). The concave surface is curved with a hydrophobic interface stretch and positively charged residues which bind to core promoter TATA DNA with nanomolar affinity¹⁸. The convex surface was shown to mediate binding with other transcription factors, with both negative and positive regulatory properties. In particular, helix2 (H2), on the convex surface is rich in basic amino acids, and extensive surface residue mutational scans in critical TBP interaction sites have clarified the global gene regulation by TBP interacting complexes¹⁹. Furthermore, the N and C lobe regions facilitate the interaction with the general transcription factors (see below). Considering the essential role of TBP in all distinct RNA Pols transcription machinery, the transcriptional output can be modulated via their interacting complex, which could both positively and negatively regulates the transcription complex assembly.

2.3. Regulation of TBP and its interplay with interactors

Abundant biochemical and few structural information are available on the several transcription factors associating with the TBP subunit in the context of general transcription factor complex, such as TBP itself forming dimer, TAF1, TFIIA, TFIIB, Mot1, NC2, Brf1, and the transcriptional activators (transactivators), including p53, Gal4, Sp1, VP16, c-Myc and Rap1. A list of structural information on TBP and in complex with transcription factors is briefly summarized in **Table 1** and will be further elaborated in later sections.

TBP regulation on itself and on the -TATA complex

TBP structures bound to TATA box revealed that the conformation of TBP’s saddle-shaped structure remains intact. However, TBP binding to DNA leads to a ~90 degree bend in the DNA, by partial unwinding of the eight base pair TATA sequence through insertion of TBPs concave hydrophobic Phe residues into the minor groove of the promoter DNA (**Fig. 3a**)¹⁸. In the absence of TATA DNA, TBP can also self-associate

to TBP homodimers, which occludes its concave surface and inhibits the DNA binding, as evidenced by the crystal structure²⁰ (**Fig. 3b**). Notably, TBP cellular concentration is functionally related to dimer formation and thereof directly regulates the expression of specific genes³³. Both *in vivo* and *in vitro* experiments consistently supports that TBP dimer formation is slow and that this process is the rate limiting step for transcription initiation³³; thus, the dimer formation is a physiologically important negative regulator of gene expression.

TBP regulation by TAFs

TBP binding to promoter DNA can be regulated by TFIID components, namely its associated factors (TFIID-TAFs). Among several TAFs, TAF1 and TAF7 are the only associated factors of TFIID to interact directly with TBP³⁴. TAF1, with homologues in all eukaryotes, including drosophila dTAF1/TAF₂₃₀, yeast yTAF1/TAF₁₄₅ and human hTAF1, is the largest and is also an integral subunit. It is functionally the most diverse TBP-associated factor, and considered to be a platform for assembly of the entire TFIID⁹. TAF1 is an essential scaffold for TFIID assembly through its interaction with TBP and other TAFs⁹. The deletion of TAF1 N-terminal Domain (TAND) region strictly affects TBP association, although the binding to other TAFs is retained²⁴. TAND domains are composed of TAND1 and TAND2, which bind to the concave and convex surface of TBP, respectively³⁵. Both *in vitro* and *in vivo* studies have supported the finding that abolishing the TAF1 interaction to TBP, functionally abrogates TFIID-mediate transcription initiation³⁶. Interestingly, dTAF1 mediates monoubiquitination of histones within the TFIID subunit, by its additional TAF1 Ubiquitin-Activating/ Conjugating activity (UBAC)³⁷, in addition to the protein kinase and histone acetyltransferase activity, thus providing a possible link to the ubiquitination-mediated activation of gene transcription³⁷.

The N-terminal region of TAF1 occludes DNA association to TBP

The solution structure of dTAF1-TAND1 region in complex with TBP shows the occlusion of the TATA box binding surface by its binding to the TAND1 domain (**Fig. 3b**)²⁴. Structural comparison of TBP-dTAF1 and TBP-TATA complexes unveils that the TBP concave surface is targeted through a distinct structural mode of binding between these two complexes²⁴, and that the N-terminal TAND region is a largely autonomous regulator of transcription, which is also functional when attached to several other TAFs in TFIID complex. The yeast TAND (yTAND) domains are also unstructured, and upon binding the TBP subunit, they forms secondary structural elements similar to dTAF1 (ref³⁸). The yTAND1 domain independently acts as an activation domain^{38,39}, whereas the yTAND2 region has an inhibitory effect on

transcription³⁹ and competes with TFIIA in binding to TBP in a dynamic interplay^{40,41}. It is intriguing to posit that the TAF1 inhibition of the TBP binding to the core promoter may possibly inhibit the partially assembled TFIID complexes from nucleating PIC complex formation and therefore avoids non-specific transcription initiation. Several biochemical studies show that inhibitory domains of TAF1 repress TBP-DNA and TBP-TFIIA complex formations, although competitively TFIIA derepresses and enhances the TBP-DNA complex. But the structural understanding of how TAF1 competes with TFIIA for the overlapping interaction surface on TBP at molecular level as well as the mechanistic interplay between the transcription factors and possible influence on transcription initiation remains to be explored.

Regulation of TFIIA association to the TBP-TATA complex

TFIIA mediates the stabilization of TBP-DNA association and regulates TBP and TFIID dimer formations, and thereafter transcription complex assembly⁹. The ternary structure of the TBP-DNA-TFIIA complex unveils that interaction on the TBP happens primarily using two anchor regions^{22,23}; (i) the most extensive anchoring by TFIIA occurs near its N-terminal lobe and H1 region of TBP, and (ii) the aromatic Tyr residue of TFIIA contacts TBP-convex surface (**Fig. 3c**). The molecular basis of TFIIA region that binds to H2 residues on TBP are not understood, as the crystal structure lacks the electron density due to residue flexibility, the exception being the aromatic Tyr residue which binds to the hydrophobic groove (**Fig. 3c**).

Regulation of TBP by RNA Pol III transcription factor Brf1

Similar to TFIID being the core initiation factor of RNA Pol II transcription, the central transcription initiation factor of the RNA Pol III transcription machinery and promoter recognition is the TFIIB, a multi-subunit complex composed of TBP, TFIIB related factor (Brf1) and Bdp1. Biochemical and structural studies have shown that Brf1 directly interacts with TBP in a similar manner as to TFIIA, with additional principal anchoring on the TBP convex surface²⁹ (**Fig. 3d**).

Table 1: Structural information on TBP and its complex with transcriptional regulators

TBP and its direct interactors	Structure reports on	PDB entry	Refs
TBP only	Molecular saddle-shaped structure of TBP predominantly self-associates into dimer	1TBP	17,20
TBP-DNA	TBP recognizes TATA promoter by bending DNA and interacting with the minor groove	1YTB	18,21
TBP-DNA-TFIIA	TFIIA stabilizes TBP and TFIID-DNA binding complexes, and also regulates the binding of transcription inhibitors	1NH2	22,23
TBP-dTAF1	Solution structure of TAND domain from dTAF1 in complex with TBP; it shows that TAND structural interface is occluded in the concave surface of TBP	1TBA	24
TBP-yTAF1	TBP-bound TAND domains from TAF1; complex structure provides the detailed molecular pattern of interactions involving both transcriptionally activating and repressing regions in TBP	4B0A	P -I 25
TBP-TFIIB	TFIIB binds TBP and recruits Pol II subunits	1VOL	26
TBP-Mot1	N-terminal region of Mot1 ‘latch’ binds to TBP concave surface; it forms a structural elements to inhibit the promoter binding	3OC3	27
TBP-DNA-NC2	NC2 heterodimer binds and stabilizes TBP-TATA complex, provided clues as to inhibition of TFIIA and TFIIB binding	1JF1	28
TBP-DNA-Brf1	Ternary complex structure reveals the flexible region of Brf1 interfaces to TBP	1NGM	29
Cooperative interplay of transcription regulators in TFIID for transcription initiation			
TBP-DNA-NC2-Mot1	Mot1 ‘latch’ region is unstructured, in presence of TBP- DNA complex	4WZS	30
TBP-TFIIA-Rap1p	Transactivator Rap1p interacting with TAF2 and recruiting TFIIA for activation	EM	31
Transactivators-TFIID complexes	Insights into the activator-dependent TFIID assembly and the modulation of activity. Transactivators such as p53, Sp1 and c-Jun are studied here.	EM	32
P -I, Paper I. EM for cryo-Electron Microscopy			

TBP and its play with interactors

Figure 3. Structure of TBP and in complex with other transcriptional regulators. (a) Center of the panel shows TBP (grey) in complex with TATA-box DNA (wheat), and shows the interaction surfaces (highlighted in grey) that contribute from various TBP complexes,

represented with their structures of TBP in complex with **(b)** –TBP on itself (green), -dTAF1 (pink) and -Mot1 (brown), **(c)** DNA-TFIIA (orange), **(d)** DNA-Brfl (magenta) and **(e)** DNA-TFIIB (sand). The structural names and the characteristic roles are annotated, and the related PDB codes are listed in **Table 1**.

TFIIB enhances TBP-TATA binding

The transcription factor, TFIIB brings the TBP-promoter DNA complex to Pol II, near its active center cleft, by forming a complex of TBP with promoter DNA and recruiting Pol II. The C-terminal B-core region of TFIIB contains two cyclin domains that interact with TBP (on the H1' and C lobe) and upstream of the TATA promoter DNA (BRE promoter element), leading to stabilization of DNA bending²⁶ (**Fig. 3e**). Thus, in addition to TFIIA, TFIIB also efficiently enhances TBP binding to the TATA promoter.

Mot1 dissociates the TBP-DNA-NC2 transcription inhibitory complex

Mot1 (Modifier of transcription 1; BTAFl in human) dissociates the TBP-DNA complex in an ATP-dependent manner and functions as a TBP chaperone. Studies from *in vivo* cellular experiments have provided evidence that TBP resides in complex with Mot1, to prevent non-specific promoter DNA interactions⁹. The N-terminal domain of Mot1 binds TBP, primarily at two regions: to the convex surface of the TBP H2 region of the N-terminal lobe and to the concave surface by the Mot1's large loop, termed as the 'latch' region which folds-on-binding upon TBP association (**Fig. 3b**)²⁷. This structural data further provides interesting clues for how Mot1 and transcription regulator, NC2 (Negative cofactor 2) synergistically, increase TBP mobility on chromatin as well as reset gene expression by jointly removing TBP from TATA elements after transcription²⁷. In support of this, the structure of NC2 with the TBP-promoter DNA complex reveals an inhibitory state²⁸, because it sterically occludes the interface for TFIIA and TFIIB association, thereafter preventing the TBP-mediated nucleation of PIC assembly⁹. However, the structural complex of Mot1-TBP-DNA-NC2 may well provide an explanation for the role of phosphorylation on the Ser residues of NC2 in modulating the binding⁴² with the TATA element. In fact, the hypothesis might be that phosphorylated NC2 creates an electrostatic repulsion to the backbone of TATA DNA, leading to derepression of TBP-TATA complex by dissociating NC2, and thus consequently relieving the global repression in transcription.

Summary for the TBP-containing complexes

Taken together, the TBP-containing complexes (**Fig. 3**) imply that TBP surfaces that are vital for transcriptional activity are: (i) the hydrophobic concave surface of TBP that binds to TATA promoter and several other inhibitory domain of transcription factors; (ii) the H2 positively charged region, which binds to TFIIA and transiently with several other transcription factors; (iii) the TBP C lobe region that binds to TFIIB, and (iv) the additional specific interactions with TAF1 or Brfl to the N-terminal lateral face of TBP. In summary, the structure-function studies have significantly implied that overlaps within TBP surfaces interacting with various complexes like that of TFIIA, NC2, TAF1, TFIIB, Brfl, BTAF1 or Mot1 are extensive. Therefore, coordinated interplay with transcription factors is necessary to achieve transcription initiation. Notably, the TBP-interacting fragments of these transcription factors binds to the TBP surfaces in a dynamic way, with few structural reorganizations, and competing for overlapping regions on TBP. Thus understanding the molecular basis of these mechanisms will shed light on the regulation of transcription initiation and on the modulation of its activity.

2.4. TFIID activity modulated by gene-specific transcription activators

In transcription regulatory mechanisms, TFIID is known to be the target of several gene-specific activators or transactivators, which play an important role in transcription initiation. Transcriptional activator proteins bind to regulatory DNA sequences in the vicinity of responsive genes by themselves or in complex with other regulatory proteins, although increasing evidence consistently supports that the transactivator domain (TAD) regulates the transcriptional initiation complex by employing the direct interaction within the TFIID transcription complex⁹.

Transactivator binding to TFIID

The 2012 Nobel Prize in Physiology or Medicine was awarded to the discovery that mature fibroblast cells can be reprogrammed into induced pluripotent stem cells, by supplementing with a set of four transcription factors: Oct4, Sox2, KLF3 and c-Myc⁴³. This further strengthens the critical role of transactivators in driving cell-specific programs by coordinately functioning in a transcriptional concert and interacting with the other basal transcription factors⁹. At much lower resolution, three-dimensional electron microscopy (EM), has captured interaction of transactivators p53, Sp1, c-Jun (ref³²) and Rap1 (ref³¹) with the TFIID multiprotein assembly, furthermore indicating the direct interaction of transactivators to TBP and TFIIA within the TFIID assembly. These studies provides interesting clues as to the functional process of DNA looping and the shape of transactivator targets binding to TFIID and mediating the activator-

dependent transcription^{31,32} (**Table 1**). The initial studies of competitive binding between the transcriptional activator and TFIID for regulation includes the acidic activator VP16 and the TAF1 competitive interactions onto TBP concave surface, unveils the functionally important TBP residues³⁵. Furthermore, the mutational analyses on the amino-terminal transactivation domain of c-Myc have provided interesting clues on the direct interaction with TBP⁴⁴ and with TFIIB, by undergoing a significant structural reorganization in the transactivation domain^{45,46}. Dynamic interactions observed within these complexes also highlight the transactivation domains binding to TFIID, modulate PIC formation through its biological dynamic assemblies and thereby activating the transcription. However, the biophysical and structural knowledge at the molecular level of how transcriptional activators regulates TFIID activity or PIC assembly is still limited, which severely hampers our understanding of gene regulatory mechanisms.

Chapter 3.

Introduction to the MarR transcriptional regulators

-With a focus on structural insights into MarR proteins

3.1. MarR, a regulator of multiple antibiotic resistances

Bacterial multi-drug resistant strains of *E. coli*, *S. aureus* and *P. aeruginosa* are amongst the major human pathogens and cause bacterial infections in the community, in particular to immunocompromised patients, and are quite challenging to treat with the current antimicrobial chemotherapy. Proteins of the MarR (Multiple antibiotic response Regulator) family regulate the transcription of operons encoding a drug efflux pump and therefore involved in antibiotic stress response (**Fig. 4a**) (reviewed in ref ⁴⁷). The founding member of this family, the MarR protein, was first identified in *Escherichia coli* and the proteins belonging to this family are distributed throughout the bacterial and the archaeal species, counting to almost 12,000 MarR-like proteins identified, but the physiological role is only known for around 100 members so far⁴⁸ (Refer to **Table 2** for the list of MarR homologues, and that are encoded in respective bacterial organisms). In general, MarR homologues function as molecular systems that are critical to bacterial physiology such as responding to specific ligands and metabolites, oxidative stress agents, virulence or pathogenic factors and multi-drug resistance. Clinical observations of multidrug resistance in bacterial phenotypes from infected patient samples have showed that the increased gain of resistance is due to MarR derepression, as a consequence, hyperexpression of the protein efflux pump, and that mainly because of spontaneous mutations in MarR transcription regulators⁴⁷. Since an understanding of the mechanism of the MarR family in regulation is critical for the efficient treatment of bacterial infections, detailed structural and functional characterization of the MarR proteins would provide better understanding of the resistance mechanisms, and may possibly help to rational design of therapeutic targets to overcome the current shortcomings in antimicrobial chemotherapy.

3.2. Structural features of MarR homologues

Members of the MarR family are poorly conserved in amino acid sequence (less than 25%), but share a common triangular shape within the family. The first crystal structure was the MarR from *E.coli* ⁴⁹ and structures of many MarR homologues have since become available (see **Table 2**). MarR homologues operate as homodimers to recognize specific symmetric DNA targets, and share a functional role to recognize a

large variety of signal molecules that regulate this binding. The structure of a MarR monomer subunit is composed of a helical dimerization domain (H1, H5 and H6) and winged (W) Helix-turn-Helix (HTH, H3-H4-S1-W-S2) DNA-binding domain, where H4 is referred as DNA-recognition helix (**Fig. 4b**). Thus the MarR family is a subgroup of winged-HTH family of DNA-binding transcription factors⁴⁷. The dimerization domain helices of both monomer subunits create a compact dimer that is stabilized by the extensive hydrophobic and hydrogen bonding interactions between the interface residues (**Fig. 4c**) (ref⁴⁹).

MarR-DNA bound complex

The crystal structures of MarR homologues bound to DNA, currently including OhrR-DNA⁵⁰, SlyA-DNA⁵¹, ST1710-DNA⁵² and MepR-DNA⁵³, have provided structural insights about the specificity underlying the recognition of precise promoter DNA in individual MarR homologues (**Table 2**). The dimerization domain and the wHTH domain are connected by the helix2, H2, which maintains the relative orientations of the subdomains in the intersubunit dimer⁴⁸. While mutational analyses on the dimerization helices have been shown to impact DNA interaction⁵⁴, the key pattern recognizing the DNA promoter region is deployed in the wHTH motif⁴⁷. Importantly, the MarR-DNA bound structures consistently shows that the H4 and H4', DNA-recognition helices from wHTH motif binds to the major groove and that the Wings from each monomer subunit places specific MarR-contacts with the minor groove of the DNA (**Fig. 4c**) (ref⁴⁸). Comparing all the DNA bound MarR structures (**Table 2**), the relative position of DNA-recognition H4 and H4' helices are placed in similar positions and the distance measure of the mid-points of these two helices (see inset in **Fig. 4d**), shows a range within the distance of 30-31 Å. Thus, MarR-DNA bound conformation in the crystal structure shows a conformation states that dictates the DNA bound MarR conformations.

Ensemble of pre-existing conformations in MarR apo crystal structures

Several crystal structures of MarR homologues in the absence of effector or DNA (herein named “apo”) consistently present a similar structural architecture (**Table 2**). Furthermore, the MexR repressor from *P. aeruginosa* have additionally provided the first glimpse of MarR proteins being considerable flexible by showing four different dimers in the crystallographic asymmetric unit (ref⁵⁵). The distance measure between the DNA-recognition helices of H4 and H4' mid-point residues (see inset in **Fig. 4d**) on the four different conformations of MexR apo dimers ranging from 23Å to 29Å, and in comparison with the aforementioned DNA-bound MarR structures reveals that

MexR apo consists of a mixture of pre-existing conformations that are both compliant and incompatible with DNA binding⁵⁵.

In agreement with the MarR apo conformations, also the MepR apo crystal structures from two independent structural studies have indicated a similar MarR structural topology, but with the different conformations that largely differ in their respective relative orientations of the inter-subunit dimer, so placing the DNA-recognition helices far apart (**Table 2**). The distance measure of the mid-point residues of DNA-recognition helices shows that one MepR apo structures presents 36Å (ref⁵⁶), closely resembling the DNA-bound structure and the other one shows the most open MarR structure of MepR H4-H4' with a distance of 43Å (ref⁵⁷). In addition, the MepR mutant F27L SeMet -substituted crystal structure, presented the most wide open conformation incompatible with DNA binding, where the distance to the mid-point of H4-H4' helices is 58Å (ref⁵⁶). Therefore, considering all the above mentioned structures of MepR apo of two different conformations^{56,57}, MepR F27L mutant⁵⁶ and MepR-DNA bound complex⁵³, these structures present a range of conformations that are present in the crystal structures of MepR and provided the ensemble of pre-existing conformations (**Fig. 4d**). In contrast with the above observations of MexR and MepR being crystallized in various dimer conformations, the HucR apo structure presents a single conformation that is more similar to the DNA bound complex with the H4-H4' helices distance of 29Å (ref⁵⁸). The similar observation is also noted in another MarR homologue, such as PcaV, which in its apo form adopts the DNA binding conformations⁵⁹.

Some MarR members undergo ligand-binding and/or redox-directed, conformational changes

Few MarR members are shown to bind small molecule ligands inhibiting its repression activity of DNA binding both *in vitro* and *in vivo*, and thereby functionally linking environmental signals with gene expression responses⁴⁸. The crystal structure of *E.coli* MarR only crystallized in presence of salicylate⁴⁹, although the physiological role of this ligand bound is not identified, but the MarR-ligand bound shows the structure incompatible with DNA binding. With the continued structural studies on MarR homologues, it is now evident that many MarR members are also regulated by ligand-binding, includes MarR, MTH313, ST1710, all bound to sodium salicylate^{49,52}, and the structures of MarR homologues bound to natural ligand, such as HucR-urate⁶⁰ and TcaR-antibiotics⁶¹ (**Table 2**). The salicylate binding site is located in a deep pocket at the dimerization and DNA binding lobe, inducing a conformational change that appears to be incompatible with DNA binding⁴⁸. While the antibiotic bound complexes have recently pointed out the essential cavities for drug binding, as

evidenced by the TcaR-drug bound structures⁶¹, further arguments underpin that the interfaces are located at the helical dimerization domain and the winged HTH DNA-recognition region. As a common feature, the MarR members that are in complex with ligand or antibiotic-bound (**Table 2**), present a single conformation structure.

In addition to the ligand bound MarR derepression, some members of MarR undergo inhibition of repressor activity by direct binding to effectors such as oxidative agents, which creates specific Cys disulfide bonds within the intersubunit dimer and restricts the associated flexibility. For instance, structures of OhrR upon oxidation have structurally formed a conformation that prevents DNA binding⁶². Of note, the phosphorylation-based regulation also occurs in some MarR member, for instance, the DNA dissociation of SarA is triggered by Cys-phosphorylation, at the same Cys residue that undergoes oxidation-dependent DNA dissociation⁶³. Therefore, the MarR homologues deploy multiple ways of regulation to modulate their target gene regulation.

Table2: Structural reports on MarR family of transcriptional regulators

MarR homologue (apo or in complex)		Bacterial organism	Structural insights and conformational states	PDB entry	Refs
MarR –salicylate (SAL) bound		<i>E. coli</i>	First identified structure of MarR members, and forms a dimer with a single conformation in crystal with two SALs bound.	1JGS	49
MexR	apo	<i>P. aeruginosa</i>	Four different conformations of dimers are observed and presents for allosteric regulation in MarR member	1LNW	55
	-ArmR peptide complex		ArmR peptide binding at the MexR inter-subunit core, stabilizes the conformation that prevents DNA binding	3ECH	64
	Oxidized state		Subtle changes in overall structure are due to Cys disulfide bridge within the inter-subunit dimer, thus obviously limits their relative orientations for the required DNA binding.	3MEX	65
	R21W		MDR mutant analysis unveil the restricted flexibility in the dimerization helices	4ZZL	P-II
OhrR	DNA	<i>B. subtilis</i>	Three different conformations observed and the flexibility resides in winged-HTH motifs	1Z9C	50
	Oxidized state	<i>X. campestris</i>	Upon specific Cys residue oxidation and disulfide bond formation, MarR undergoes the effector-induced derepression.	2PFB	62
Continued to next page					

Table 2. Continued

MarR homologue (apo or in complex)		Bacterial organism	Structural insights and conformational states	PDB entry	Refs
SlyA-	apo	<i>S. typhimurium</i>	The flexibility and specificity that co-exist in winged HTH domain revealed the basis for DNA recognition	3QPT 3Q5F	51
	DNA complex				
ST1710	SAL ligand	<i>S. tokodaii</i>	Critical DNA binding residues were identified and two SAL ligand binding pockets agrees with the other MarR homologues	3GF2 3GFI	52
	DNA				
MTH313 apo and -SAL bound		<i>M. thermoauto trophicum</i>	Single conformation observed for both apo and ligand bound, and the studies revealed the conformational changes that are resulted from ligand bound	3BPX	66
TcaR –Antibiotics		<i>S. cerevisiae</i>	Structurally explains the MarR regulation for its associated plasticity in multidrug binding pocket. Antibiotic binding locates at the dimer interface, thus binding causes the conformational changes in winged HTH motif distances and inhibits DNA association.	4EJU 4EJV 4EJW	61
HucR		<i>D. radiodurans</i>	Single conformation dimer was observed for HucR apo structure that resembles the DNA binding conformation.	2FBK	58
PcaV - ligand bound		<i>S. coelicolor</i>	PcaV apo structure and ligand bound, reveals the Helix 1 residues, in particular, the Arg residue interacting with the anionic ligand molecule and that conserved in other MarR homologues	4G9Y	59
MepR	apo	<i>S. aureus</i>	Two independent structural studies reveal distinct dimers of single conformations.	3ECO 4L9J	53, 56, 57
	DNA complex		Three different conformations	4LLN	
	Single mutant-Q18P, F27L and A103V		MepR mutant crystal structures present its local changes, and those inhibit the DNA association	4L9T 4L9V	
P-II. Paper –II (this study)					

3.3. MexR, a MarR repressor of the *P. aeruginosa* efflux pumps

P. aeruginosa is an opportunistic well-known human pathogen as it causes a broad spectrum of severe infections in immunocompromised individuals, and in particular, in patients with cystic fibrosis, with potentially life-threatening complications in some individuals (reviewed in ref ⁶⁷). The lists of infections that are associated with this bacterium are incredibly high and treating the infections is extremely difficult not only due to its innate resistance to various antimicrobial agents, but also to the possibility of acquiring additional higher levels of multidrug resistance⁶⁸. In fact, its intrinsic biological properties, including the low permeability of its cell wall, genetic variation to express efflux systems and mutation in chromosomal genes, conveys intrinsic resistance to many diverse antimicrobial agents. In addition, among a wide range of resistance mechanisms, the presence of efflux systems is particularly well developed in *P. aeruginosa*. Efflux pump systems are three-component protein assemblies that actively transfer toxic substances across the cell wall and outside the bacterium⁴⁸. So far, seven RND (Resistance, Nodulation and cell Division) family of multidrug efflux systems have been identified⁶⁹ and clearly there have been major interests in clinical perspectives to target this pump with therapeutic agents, to inhibit the efflux pump activity of extruding multidrug⁷⁰. Therefore, an in-depth analysis at a molecular level of how this efflux pump can be regulated is of special interest and may also serve as a model system to understand the structure-function of multidrug resistance in human pathogens.

The principal efflux pump that contributes to the multidrug resistance in *P. aeruginosa* is the MexAB-OprM pump, which possesses a particularly wide range of substrate binding and is encoded by the *mexAB-oprM* operon⁷¹. Studies on this pump have revealed that *mexAB-oprM* gene expression is regulated by three regulator genes: *mexR*, *nalC* and *nalD*. The repressor, MexR is a negative regulator of this efflux pump system (**Fig. 4a**), as it binds as a dimer to two promoter regions, to represses both the expression of the *mexR* and *mexAB-oprM* genes and belongs to the MarR superfamily of bacterial transcription regulators⁷². Likewise to the MexR direct binding to the promoter region, the NalD binds in proximity of the promoter site directly to represses the pump expression as well. The last regulator protein, NalC, is a repressor of the *armR* gene expression and indirectly regulates the efflux pump expression. However, ArmR can interact with MexR and the overexpression of ArmR leads to MexR derepression through direct binding of the two proteins and consequent release of MexR from DNA⁶⁴; this leads to the expression of the efflux pump and to the development of multidrug resistance.

Figure 4. Structure of MarR family of proteins bound to DNA and ensemble of conformations observed in the MarR apo crystal structures. (a) MexR represses the

expression of *mexAB-oprM* and *mexR* genes, and upon receiving the reactive signals, MexR derepression occurs to activate the transcriptional expression of the efflux pump (b) MexR apo monomer subunit colored in grey and secondary structural elements are annotated. (c) Structural model of MexR homodimer (subunits in grey and black) bound to DNA (ribbon, white), modeled based on the superposition of the dimerization region (residues 5-25 & 108-139) of MexR apo structure (only C α atoms) onto MepR-DNA complex. The DNA recognition helices are marked. (d) Superposition of crystal structures of MepR-DNA (one subunit in cartoon and the other in ribbon representation except the recognition helices), two conformations of MepR apo (ribbon), and MepR mutant F27L (ribbon)⁵⁶ in dimerization region (using only C α atoms) to present the associated structural heterogeneity and each of the DNA-recognition helices (colored black, cartoon) are showed. The structural regions of DNA, some regions of winged-HTH motif are removed for clarity purposes. The inset shows only H4 and H4' helices (represented in box), illustrating the relative movements of H4' in various conformations, midpoints and arrows indicated to show the center. For PDB codes and references, see **Table 2**.

Multidrug resistant MexR variants disengages DNA binding and the structural analysis reveals the possible ligand-binding pocket

The majority of the clinical cases of hyperexpression in the *mexAB-oprM* pump is due to mutations in the three regulatory genes: *nalB* (*mexR*), *nalC* and *nalD*. Several reports have consistently showed that the increasing clinical cases of multidrug resistance occur due to mutations in *mexR* gene; in particular, there is an emerging resistance to β -Lactam antibiotics⁷³. The characteristic analysis of clinical specimens of *P. aeruginosa* overexpressing the MexAB-oprM pump, presented the list of spontaneous point mutations on MexR, that leads to MexR derepression of the operon⁷³. Extensive biophysical analysis on the wild-type MexR show that mutations at the interface for dimer formation affect stability but with maintained structure⁵⁴. Some mutants are referred as 'allosteric' or 'distant' mutants, which confer the activity of incompatible DNA binding through mutations that are distant from DNA binding and dimerization helices. Similar clinical relevant mutants are also observed in other MarR members, includes MarR⁷⁴ and MepR⁵⁶. The crystal structures of MepR apo form and the clinically-linked mutants (**Table 2**) have provided the initial clues for allosteric regulation by presenting the effect on its local changes, that relatively alters the position of DNA-recognition helices in unfavorable orientations to prevent the DNA binding⁵⁶. The MepR point mutants, including residue Q18P, F27L and A103V, although not involved in dimer interfaces, are located in the dimerization helix and involved in intramolecular contacts, and therefore referred as allosteric mutants as well⁵⁶.

The most recent structural characterization of PcaV apo and -ligand bound crystal structures (**Table 2**), presented the residue Arg15 in Helix1 as a ‘gatekeeper’ residue, since mutant Arg15Ala completely abolishes the DNA binding by disrupting its structural network⁵⁹. Therefore, ligand binding at this region would favorably lead to disrupt the repressor-DNA binding, as evidenced by the structure of PcaV-ligand bound⁵⁹. In agreement with the ligand binding pocket at Helix1, the residue Arg16 from MTH313 (MarR homologue) is involved in salicylate binding, as evidenced by structure of MTH313-salicylate bound complex⁶⁶. The extensive mutational analysis of this ligand binding pocket have provided in-depth details of interface residues, especially the side-chain of a basic residue located on Helix 1, with the carboxylate group of salicylate⁷⁵. Similarly, the structure of another MarR homologue, ST1310-salicylate bound complex further evident that the basic residue of Helix 1, Arg20 from ST1310, serves the same interface and binding mode⁵². Furthermore, the structural comparison to other MarR homologue such as TcaR, SlyA and ST1310 (**Table 2**) shows the potential to bind the structurally related anionic ligands at the same site through an adaptable ligand binding pocket⁷⁵. Although the local changes upon mutations are studied, the distant structural changes that consequently affects the DNA-recognition helices causing the incompatible DNA binding remains to be explored.

Regulation of MexR by anti-repressor effectors and oxidation

Interestingly, MexR is so far the only MarR family member that has been shown to be modulated by anti-repressor peptide⁶⁴, although other MarR homologue are regulated by small hydrophobic or phenolic compounds⁴⁸. The crystal structures of MexR in complex with ArmR peptide of 26amino acids have presents an extensive binding interface that is mainly hydrophobic⁶⁴. An interesting observation is that, upon ArmR binding, the MexR undergoes more symmetrical dimer compared to MexR apo structures⁵⁵. In the context of the DNA-recognition helices, upon ArmR interaction to MexR, the helices H4-H4’ move upwards to 13Å and the wings relatively positions at a distance of 18Å, which appears to be incompatible with DNA binding⁶⁴. Similarly, the distance between the DNA-recognition helices H4 and H4’, have revealed a distance of around 25Å suggesting unfavorable DNA binding.

Likewise other MarR members, MexR also undergoes derepression of promoter DNA by oxidative stress conditions⁷⁶. The crystal structure of the oxidized MexR unveil the structural changes⁶⁵ that are evident by comparing it with the MexR apo structures, and provides the evidence that the cysteine disulfide bond formation within the MexR dimer would simply inhibit a MexR-DNA complex association because of its steric clash with the DNA backbone. Unlike many different

conformations observed in the crystal structure of MexR apo, the MexR oxidized state structure presents a single conformation⁶⁵. Although the disulfide bond restricts the binding to DNA by inhibiting the interdomain flexibility, further understanding of how the DNA-recognition helices distance is maintained, needs more investigation. To conclude, despite the crystal structure observations of structural heterogeneity within the MarR apo states, the mechanism underlying the conformational states in MarR homologues remains to be explored. The potential to shed light into these questions may possibly unveil the region or part of structures, that may well be a major interests in clinical perspectives to target this DNA repressor with rationally designed therapeutic agents.

Chapter 4.

Introduction to protein ubiquitination

-With a focus on structural insights into the E2 and E3 (RING-type) ubiquitin-binding proteins and the TRIM E3 ligases

Post-translational modifications on target proteins (substrate) increase the functional diversity of the substrate, and they are commonly employed to alter the cellular fate of that target. One such post-translational modification is tagging with the small protein ubiquitin (Ub) molecule, which is used in eukaryotic cells as an important signaling module. This ubiquitination stage in proteins is defined as ubiquitination (also referred as ubiquitylation), and is one of the most dynamic and highly regulated steps in eukaryotes. Numerous cellular processes depend on it, including intracellular protein turnover, cell-cycle control, transcriptional regulation, growth factor signaling, DNA repair, gene silencing and stress response^{77,78}. Since ubiquitination regulates from cell proliferation to programmed cell death, any defects in the functional regulation of ubiquitination can lead to cellular disorders, which have been implicated in a wide range of diseases, such as cancers and neurodegenerative diseases^{79,80}. To highlight the importance of ubiquitination, the 2004 Nobel Prize in Chemistry was awarded to Aaron Ciechanover, Avram Hershko, and Irwin Rose for their initial discovery in ubiquitin-mediated proteasomal degradation of intracellular proteins⁸¹ and for defining the characteristic feature in the ubiquitination system.

4.1. The ubiquitination system and ubiquitin-binding proteins

The covalent attachment of Ub to targeted proteins is a multistep orchestrated action, involving a series of three sequential enzymes: the Ub-activating Enzymes (E1), the Ub-conjugating enzymes (E2) and the Ub-ligase enzyme (E3) (**Fig. 5a**). The process is initiated by the ATP-dependent activation of Ub by E1, where a thioester conjugate is formed between the carboxyl-terminal Gly76 of Ub and the active site Cys residue of E1. Eventually, the Ub-loaded E1 interacts with the E2 enzyme and transfers the Ub to the active Cys of the E2 via a similar thioester conjugate, thus forming the E2~Ub ('~' refers to thioester linkage and donor Ub) formation. The E3 ligase completes the transfer of Ub from an activated E2~Ub to the targeted substrate, typically resulting in the covalent attachment of the carboxyl-terminal Gly76 of Ub to

ϵ -amino group of site-specific Lys residues, or to the extreme N-terminus, of the substrate, which is destined to be ubiquitinated⁷⁸ (**Fig. 5a**). The two most relevant features in ubiquitination are the number of Ub molecules attached and the position of the Ub linkages, which together determine the fate of the ubiquitinated substrate.

In humans, there are 2 E1s, ~40 E2s and ~617 E3s, which attach the Ub molecules to thousands of cellular substrates⁸². Remarkably, the number of human genes encoding the E3 ligases (~617 E3s) exceeds the number of genes for protein kinases (~ 518 genes encodes for the kinases), thus hinting at a high number of intracellular proteins that undergo ubiquitination⁸². The limited number of E2s suggests that a single E2 may interact with many E3 ligases. The majority of E3 ligases bind the substrates with distinct substrate-binding motif, and thereby determine the substrate specificity of the pathway, but also help to transfer the activated Ub from the E2s to substrates (recently reviewed in ref^{83,84}). Thus E3 ligases serve dual roles as both ‘molecular-matchmaker’ (the definition coined in review ref⁸⁴) and catalyst. E3 ligases are classified into three categories based on the functional mechanism through which Ub is transferred from E2 to the substrate: (i) the homologous to E6AP carboxyl terminus (HECT), (ii) the Really Interesting New Gene (RING) and U-box family of proteins, and (iii) the RING-between-RING (RBR) family of E3 ligases^{84,85}. The HECT and RBR-type of E3 ligases form an E3~Ub intermediate step (**Fig. 5a**, see inset), prior to attaching the Ub onto a substrate lysine residue. However, the largest family of E3s is the RING-type of E3 Ub-ligases, where the E3 simultaneously binds to both E2~Ub thioester complex and substrate, through their RING and substrate-binding domain, respectively, which then mediates the direct transfer of the Ub from E2 to the lysine residue of the target substrate via an iso-peptide bond.

Many of the steps in the complex formation between ubiquitin, the enzymes (E1, E2 and E3) and the substrates use transient and overlapping interfaces (**Fig. 5c**). This is a fascinating feature and the hidden beauty in this system, which may be required to achieve maximal efficiency in the ubiquitination process. In particular, on E1 or E3 binding, the E2 structure does not appear to undergo significant conformational changes⁷⁸ and structural studies on these enzymes⁸⁶⁻⁸⁸ have provided evidence that E2 binding interfaces mutually overlap for both E1 and E3 binding. Competitive binding experiments show that E1 binds E2, to attach the Ub molecule, and subsequently E1 binding is disengaged, thus allowing E3 ligases to bind E2 on the same surface for the catalytic Ub transfer. In this way, E2 is recycled for progressive binding of the subsequent Ub in the ubiquitination process⁸⁹.

Types of Ub modifications

As mentioned earlier, the type of modifications with Ub on the target substrate dictates its cellular fate. Ub has eight potential acceptor sites: Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 and Lys63, and the amino-terminus Met1 (**Fig. 5b**), together serving as attachment sites for Ub chain assembly and for the formation of different Ub linkages⁹⁰. Monoubiquitination of a substrate often signals to altered subcellular localization and thereby alters the protein activity. Multiple monoubiquitination is also possible on a single substrate but at various lysine positions. Polyubiquitination of a substrate can result in the possibility of seven types of homotypic polyubiquitin chain linkages and each type of modification possesses different functions. The most well-characterized polyUb linkages on substrates are the Lys48-linked and Lys63-linked Ub chain linkages. Lys48-linked polyubiquitin chains represent a signal for degradation of the Ub tagged substrates by the 26S proteasome machinery, while the Lys-63 linked chains signals processes involving DNA repair, protein sorting and mediate protein-protein interactions. Moreover, mixed polyubiquitin chains are also possible, leading to the formation of multiple heterotypic Ub chain linkages on the same substrates⁹⁰.

There are also other possible types of conjugations that concern the ubiquitin-like (Ubl) modifiers to target substrate proteins. Although Ubl modifiers possess a distinct amino-acid sequence compared to ubiquitin, the overall structural fold and the mode of conjugation to the target proteins are similar. Ubl proteins include SUMO, NEDD8, FAT10 and ISG15, and the corresponding modifications, called Sumoylation, Neddylation, Fatylation and ISGylation, respectively, which all play critical roles in cancer-relevant pathways⁸⁰. For example, recent findings have enlightened a sophisticated cross-talk between different types of such modifications on the tumor suppressor p53 (ref⁹¹).

A central goal of ubiquitin structural biology is to understand the detailed molecular mechanisms underlying these complex biological processes. Given this, one could possibly design molecules that would alter the ubiquitination processes for disease treatments. The following parts illustrate the recent structural findings in the E2-E3 RING-type complex and the molecular mechanism underlying in the Ub transfer onto the substrate and how additional regulatory domains/proteins influence the Ub ligation to the site-specific residue of substrate. Finally, the knowledge of these structural findings is applied to provide mechanistic clues for the TRIM family of RING E3 ligases, in particular, TRIM21 E3 ligase potentially involving roles in mediating substrate ubiquitination.

4.2. Structural features of Ub - ‘A molecular-tagger’

Ub is an 8.5 kDa highly stable protein, comprising 76 amino acids and is evolutionary conserved from yeast to human. The structural fold of Ub adopts a compact β -grasp fold formed by a five-stranded β -sheet (β 1- β 5), an α -helix (α 1) and a 3_{10} helix (**Fig. 5b**). Turning outward from the canonical ubiquitin fold, an exposed C-terminal flexible tail, composed of residues (76 GGRLRL 71), is essential for the conjugation of Ub molecule (**Fig. 5b**). Many Ub-binding domains (UBD) have been shown to interact with a solvent exposed hydrophobic surface on the Ub molecule. The most characterized Ile44-centered hydrophobic patch includes residues Leu8, Ile44, His68, Leu69 and Val70, which are located on the β -sheet (**Fig. 5b**). Structural studies of the dynamic nature of Ub on nanosecond to microsecond timescales have provided in-depth details of the ensemble of Ub conformations and the associated flexibility required for interacting with various UBDs, as evidenced by NMR residual dipolar couplings measurements⁹⁴. The observed flexible regions are located on the loop region between β 1 and β 2, and the loop residues between α 1 and β 3 (**Fig. 5b**)⁹⁴. Interestingly, these residues from the loop and the bolstering residues from the C-terminal tail contribute to another hydrophobic surface, which is Ile36-centered hydrophobic patch, including the residues Leu8, Thr9, Glu34-Pro37, Leu71 and Leu73 (**Fig. 5b**). These regions facilitate the interaction with the Lys11 linked Ub chains and they are important also for the recognition with the UBDs, DUBs and E3 ligases. Taken together, the hydrophobic surface patches of Ile44 and Ile36 of Ub form a hotspot for Ub recognition and primarily mediate interactions.

Table 3. Family of E2 enzymes

Four classes	E2 enzymes*
I , UBC only	UBE2D1-4 (UbcH5a-c), UBE2W, UBE2V2 (MMS2), UBE2N (Ubc13), UBE2L3 (UbcH7)
II , UBC + C-terminal extension	UBE2R1-2 (CDC34), UBE2S, UBE2J1-2, UBE2U
III , UBC + N-terminal extension	UBE2E1-3 (UbcH6-8), UEV-1A, UbcH10
IV , UBC + N & C-terminal extension	UBE2O, BIRC6

*Alias in parenthesis

Figure 5. The ubiquitin system, and the structural features of Ub and E2s UBC domains. (a) Schematic view of ‘The ubiquitination pathway’, where E1, E2, E3 enzyme

ubiquitylating substrate in an ATP-dependent manner, and the number of genes encoding in humans are annotated here and the inset is a lists of the functional role of the ubiquitin tagged substrate. **(b)** Structure of ubiquitin (PDB: 1UBQ, orange) with its potential eight acceptor sites (7 Lys and 1 Met, shown in sticks), C-terminal carboxyl tail residues, and the Ile36 (blue) and Ile44 (white)-containing hydrophobic patch are annotated and shown in sticks. **(c)** Structural feature of E2 enzyme, UBE2E1_{core} (PDB code: 5A4S, this study) shown in grey and the critical residues for catalytic and structural role are defined and annotated in sticks, with the active site Cys in spheres (yellow). The lined map provides the overview of interface region with different E2-interacting partners. **(d)** E2 enzyme that are both covalently conjugated and non-covalent interaction with ubiquitin adopts various conformations of E2~Ub (PDB codes: For backside Ub -2FUH, 4R62; covalent linked Ub -2GMI, 3A33; and Ub binding to helix2 interface -4V3L), highlighted are the open and closed E2~Ub conformations and also non-covalent backside interactions: canonical backside⁹² and non-canonical backside⁹³.

4.3. E2 ubiquitin-conjugating enzyme -‘the molecular choreographer’ for ubiquitination

E2 enzymes play a central role in the ubiquitin conjugation pathway and in particular in the Ub chain assembly (reviewed in ref ⁹⁵). E2s are the ‘molecular choreographer’ in the ubiquitination enzymatic pathway (as termed in ref ⁹⁶), as they possess a wide range of functions and interact with many different ubiquitin players (**Fig. 5c**). The E2 enzymes are classified into four classes based on presence and nature of flexible N- and/or C-terminal extensions, apart from the UBC domain (Ubiquitin conjugating, ref⁹⁶) core domain. The E2 classes I, -II, -III and -IV represent the UBC domain only, with the C-terminal extension, with the N-terminal extension and with both the N, C-terminal extensions, respectively (see review, ref ⁹⁷) (**Table 3**). The catalytic UBC core domain of E2s is highly conserved evolutionarily with a prolate-ellipsoid shaped comprising four α -helices (Helix 1-4), a 3_{10} helix, a four stranded anti-parallel β -sheets (β -sheet 1-4), and an active-site cysteine residue (Cys131 in UBE2E1) (**Fig.5c**). Structural and biochemical characterization of E2 enzymes have identified many crucial residues and motifs that are important for structural and catalytic function. All the E2 enzymes contain an active-site catalytic Cys residue, which forms a thioester conjugate with the C-terminal tail of Ub, thus forming the E2~Ub conjugate. Other highly conserved residues on E2 is the three-residue HPN or HCN (His-Pro/Cys-Asn) motif (**Fig. 5c**), located at the close vicinity of active site⁹⁸. Functionally, the side-chain NH₂ of the Asn from HPN/HCN motif was proposed as catalytically important in stabilizing an oxyanion intermediate during the nucleophilic attack by the incoming lysine residue on either a substrate or on the growing Ub chain⁹⁸. Recent

analysis underpins the structural role of this Asn for E2 catalytic activity but reveals that its primary involvement concerns maintaining the structure of the active site loop⁹⁹. Moreover, the Asn is further stabilized by the His residue through H-bonding within the HPN/HCN motif and thus plays an important role in structural stability as well¹⁰⁰. Altogether, the HPN/HCN motif, through its extensive network of H-bonding, is important for structural stability of the active site loop and thereby supports the E2 enzymatic activity.

E2s backside and C-terminal helix binding to E2-interacting domains regulates the activity. Several biochemical and structural studies have also intriguingly exemplified that E2 enzymes are regulated by a number of ways, and thereby employ the regulation within an individual ubiquitination system^{83,84}. Backside (E2s β -sheet region, **Fig. 5c**) interactions by non-catalytic ubiquitin E2 variant or ubiquitin or secondary E2 binding domain apart from the RING domain of E3 ligase for altering the enzymatic activity, such as R6BD from Rad18, G2BR from gp78 (**Fig. 6a**) and U7BR from Cue1p E3 ligases are studied for its respective function (reviewed^{83,84}). Further, co-activators of E3 (co-E3) or a region from substrate binding to C-terminal helices of E2 mediate site-specific ubiquitination on substrate were also identified (**Fig. 6a**).

E2 regulation by Deubiquitinase (DUBs)

While the Lys48 linked polyubiquitin substrate undergoes irreversible protein degradation through targeting to the 26S proteasome machinery, the ubiquitination of modified Ub-substrate is a modification that can be reversed by a group of peptidases de-ubiquitylases (DUBs) enzymes. DUBs remove the iso-peptide linked ubiquitin from substrates and also disassemble the polyubiquitin chains, leading to free ubiquitin that can be recycled for new substrate modifications (reviewed in ref¹⁰¹). The human genome encodes about 90 DUBs¹⁰², which exhibit the function opposite to E3 ligases and thus provide an additional specific regulation. A proteomic and mass-spectrometry study provided evidence that DUBs form complexes with E2 enzymes and E3 ligases¹⁰³, emphasizing that the DUBs can additionally regulate the specific substrate by altering its associated pairs.

Structural studies on DUB activity shed light on how DUBs form complexes with the E2 and prevent the Ub transfer, as well as the regulation of DUBs activity on itself (recent review ref¹⁰⁴). For example, the structure of the DUB protein OTUB1 with the Ub-loaded E2, Ubc13 (ref¹⁰⁵) and Ubch5 (ref^{106,107}), emphasizes that the DUBs bind at the same interface as E1 and E3, repressing the activity of E2 enzymes. Interestingly, the N-terminal residues of OTUB1 undergo structural reorganization

from a disordered state to a structured α -helix upon interaction with Ub¹⁰⁷, presenting an ideal example for energetically favorable folding upon binding and complex formation. Still, DUB regulation *in vivo* and the molecular basis of DUBs in action together with the polyubiquitinated substrate and the role of E2 on it remain to be explored. From a therapeutic perspective, understanding the DUBs role in individual and specific E2-E3 pair may possibly lead to the discovery of attractive drug target candidates.

E2s are regulated by the flexible extensions that are apart from UBC domain

The flexible extensions that turn outward from the UBC core domain are essential for some E2s, where they direct subcellular localization¹⁰⁸, stabilize the E1 interaction, act on substrate selectivity¹⁰⁸, and regulate the Ub chain assembly¹⁰⁹. As mentioned earlier, the presence of these extensions determine the classification of the E2 enzymes. The N-terminal extension in UBE2E1 (UbcH6 – class III E2, **Table 3**) inhibits the polyubiquitin chain assembly, as shown by two independent studies but with distinct mechanisms. The initial studies show that the UBE2E1 N-terminal extension regulates the Ub chain assembly on its E3 ligases¹⁰⁹, and that the N-terminal extensions are intrinsically disordered regions, thereby auto-inhibiting the Ub chain assembly formation on the cIAP2 E3 ligase¹⁰⁹. On the contrary, more recent findings using mutational analysis on UBE2E1 reveal that the reduction of ubiquitination-mediated activity towards E3/substrate is likely due to auto-ubiquitination (self-ubiquitination) on the Lys residues in the N-terminal extension¹¹⁰. Interestingly, a UbcH5a construct fused with the UBE2E1 N-terminal residues leads to similar reduction in the E3/substrate ubiquitination activity. In particular, the E2s belonging to the N-terminal extension of UBE2E family employs the regulation of Ub chain assembly formation, but the structural characterizations of the N-terminal extension and its cellular roles remain to be explored.

4.4. RING-containing E3 ligases

Structure of the RING domain

The RING structure is stabilized as a result of Zn^{2+} -binding by eight coordinating Cys/His residues arranged in a cross-braced manner with conserved spacing between the residues⁸⁵ (see also **chapter 1**, section 1.2). An exception to this rule is the U-box domains, which do not coordinate Zn^{2+} but instead comprise a frame of stabilizing hydrogen bonding networks and electrostatic interactions in the corresponding regions of the structure⁸⁵. Both RING and U-box domains contain a central helix flanked by two surface exposed loops, together with an antiparallel β -sheet, which contain highly conserved hydrophobic residues essential to pack the core tightly and provide stability to the molecule. Some RING E3 ligases also consists of an elongated N- or C-terminal helix that are outward from RING domain which is shown essential for dimerization and to support Ub transfer¹¹¹.

The relevance of oligomer state to RING function

RING type E3 ligases can be either monomeric or dimeric, or can form higher order oligomers. Examples of homodimeric RING domains include RNF4, BIRC7, cIAP, RAG1, and the U-box protein CHIP, where both the RING domains have the capacity to functionally interact with the E2 (ref⁸³). In contrast the heterodimeric RING E3's such as, Mdm2-MdmX and BRCA1-BARD1, only one RING domain appears to be functionally active to bind E2 (ref⁸³). The substrate for the Mdm2-MdmX complex, p53, is active in the tetrameric state¹¹². This raises the possibility to have four RING heterodimers surrounding the substrate to facilitate ubiquitination. In another example, the N-terminal domain of the RNF4 was shown to bind poly SUMO (small ubiquitin like modifying enzymes) substrates and recent findings have shown that, at concentrations close to *in vivo* levels, RNF4 stays predominantly monomeric and inactive. However, RNF4 is activated by dimerization in the presence of SUMO substrates resulting in auto- ubiquitination and mediate substrate ubiquitination¹¹³. Some RING type E3's such as gp78 and Cbl, have been shown to form dimers or oligomers through domains which are structurally different from that of the RING⁸³.

4.5. E2 - RING E3 interface

Crystal structure interface provides insights into E2-E3 selectivity

Recent experimental studies have enlightened that the specificity of ubiquitination on E2 by individual E3 ligases has dual purposes: to perform site-specific ubiquitination and to define the topology of Ub chains on specific substrate^{114,115}. To this end, understanding the structural basis of an individual E2-E3 complex towards mediating target ubiquitination is a major focus of structural biology in the ubiquitin field⁸⁵. The first E2- RING E3 crystal structure complex of UbcH7-c-CBL E3 revealed an interface of binding⁸⁶, that has since been observed in the dozen of E2-E3 structural complex available (**Table 4, I**). In this generic interface a similar amino-acid composition involves the Zn²⁺ bound loops from RING domain that interacts with the canonical N-terminal Helix1, loop4 and loop7 of E2's UBC core domain, and the interface buries a total surface area of ~500 - 700 Å² (**Fig. 6a**) (ref⁹⁷). The residues from E2s, such as Phe from Loop4 'PxFPP' motif and a three-residue 'SPA' (Ser-Pro-Ala) motif from Loop7 are essential for the binding to E3 (ref⁹⁷). More specifically, from the RING interface residues, hydrophobic residues such as Val559 in cIAP2, Ile26 in BRCA1 and Trp408 in c-Cbl, have essential roles in the interfaces for the E2 Loop 4 region, and altering this region impairs the interaction.

However, despite mutual sequence conservation of E2 hydrophobic residues in the E2-E3 interface⁹⁷, many findings report the exclusive E2-E3 specificity: for example, the specific E2-E3 pair of Ube2g2-gp78 only exerts its functional role, whereas other E2s pair with gp78 E3 ligase inhibits its functional ERAD activity¹¹⁶; similar situation for Rad6-Rad18 (ref¹¹⁷) and the subsets of E2- BRCA1/BARD1 (ref¹¹⁸) and recently also for the Ube2T interactions with FANCL¹¹⁴. This implies the additional intermolecular contacts between the E2 and E3 are essential for selective interaction, and even very slight sequence variations in E2 interface residues could account for specificity of interaction with particular E3 ligases¹¹⁹. Furthermore, two independent studies using combined yeast two-hybrid screens and homology modeling evidenced several E2-E3 selective pairs and suggested that a subtle but distinct interaction network confers specificity in E2-E3 interaction^{120,121}. Notably, a significant limitation to our understanding of E2-E3 interactions is imposed by structures of RING-E2 complexes only being available for certain E2s, including UBE2D (UbcH5 family), UBE2N and UBE2L3 (**Table 4, I**) and largely absent for the class III E2s, which is known to specifically interact with TRIM family of proteins (discussed in section 4.8), the largest subfamily of the RING-type E3 ligases¹²².

a. E2 interfaces to RING, and additional non-RING & co-E3 domains

b. RING E3 ligases use a mode of Ub recognition and activation

Figure 6. Crystal structure complexes of E2-E3, E2-coE3 and ternary E2~Ub-E3. All E2s and RING domain from their respective E3 ligase are annotated. (a) Cartoon

representation of E2-RING and E2-coE3 (Pex22p). **(b)** Cartoon representation of E2~Ub-E3 complexes, the Ub and Ubl recognition residues from RING domain are shown in spheres (yellow). See **Table 4** for PDB codes and refs.

RING-type E3s indirect effect of E2 binding

In the crystal structures of all the available E2- RING E3 complex structures (**Table 4, I**), the distance between any atom from RING domain residue to the E2's active site is almost $\sim 15\text{\AA}$ (**Fig. 6a**), thus implying that RING E3s do not directly participate at the E2 active site in Ub transfer. Very little conformational change has been observed in E2 or RING domain structures upon complex formation. Some studies highlight that some E2s undergo an E3-independent activity in Ub transfer, for example E2 Cdc34 (ref¹²³), and that the presence of the RING E3s greatly enhances catalytic activity¹²⁴. Due to the lack of conformational changes on the E2-E3 binding and the distance between the E3 and the E2 active site, the Ub transfer mechanism was initially proposed as being allosterically activated. While this allosteric mechanism is not immediately evident from the available E2-RING structures, bioinformatics tool and statistical coupling analysis combined with mutational analysis have provided initial glimpses of the allosteric mechanism. In particular, residue Ile37 in UbcH5b, located at the backside of the UBC domain, has been functionally linked to catalyze through long-range communication with the E2 active site mediated by RING E3 binding¹²⁵. Furthermore, NMR Chemical Shift Perturbation (CSP) analyses show that Ile37 and the 3₁₀ helix of UbcH5b are perturbed on interaction with the CNOT4N RING domain¹²⁵. Additionally, helix2 of UbcH5b is perturbed by the CNOT4 RING, suggesting this helix may give specificity to the interaction¹²⁶. In agreement with this finding, several solution studies have shown chemical shift changes on residues distant from both the active site and the RING E3 binding interface, including E2 binding analyses with the heterodimeric BRCA1/BARD1 RING¹¹⁸ and monomeric E4B U-box¹²⁷. Taken together, while the solution studies consistently show E2-E3 binding interfaces similar to the interfaces observed in crystal structures, there are additional significant chemical shift perturbations that may provide clues to RING-mediated catalysis.

Table 4: Structural findings in RING-containing E3 ligases that reveal the catalytic mechanism of Ub transfer from E2 enzyme to site-specific substrate.

Ubiquitination complexes	R ^ε	Structural and molecular insights	PDB entry	Refs
I. Structures of E2-E3 (RING domain) complexes (Fig. 6a)				
UbcH7-c-Cbl	R	First structure of E2-RING E3 ligase reveal the interfaces	1FBV	86
UbcH5b-CNOT4	R	E2-E3 interface identified by NMR CSPs and docking	1UR6	126
UbcH5a-BRCA1/BARD1	R2'	A subset of E2s interaction with BRCA1, and comparative analysis provides a molecular basis for E2 selectivity	-	118
UbcH5b-cIAP2	R2	Subtle changes in the E2-E3 interface determines the selectivity	3EB6	139
UbcH5c-Bmi1/RING1B	R2'	Initial insights of Arg residue from RING contacting to E2 loop7 and modeled for substrate nucleosome interactions	3RPG	140
UBE2D-IDOL	R2	Critical intermolecular interfaces of IDOL to E2 provides the molecular basis for selectivity in this pair	2YHO	141
UbcH5c-E4B (U-box)	U	Solution study identifies additional interactions that are not observed in crystal interfaces of E2-E3, hence proposed for an allosteric mechanism of the enzymatic activity.	3LIZ	127
UbcH5a -RNF146 (isoADPr ligand bound)	R	Regulation of E3 activity by the ligand iso-ADP ribose binds at the interface of the RING and WWE domains.	4QPL	134
Ube2T-FANCL	R	Extensive interactions observed between additional residues of flanking helix outwards from the RING domain to E2, which provides selectivity in this pair.	4CCG	114
UbcH5c - RING1B/PCG5	R2'	Glu77 residue from RING contacts to Ub Ile-36 patch	4S3O	142
Ubc13-TRIM5	R2	RING dimerization favors Lys63-polyubiquitin chain	4TKP	143
Ubc13/Mms2-RNF8	R2	Coiled-coil domain facilitates the RING dimerization	4ORH	124
Ubc4-Not4	R	Specificity & intra-molecular H-bond of Ubc4 is identified	5AIE	144
UbcH6/UBE2E1 - TRIM21	R2	Acidic residues of TRIM E3 ligase contacts the Ub Ile-36, and a molecular basis for selectivity and pseudo-substrate	5A4S	P-IV
II. Structures of E2~Ub-E3 (RING) ternary complex (Fig. 6b)				
UbcH5c~Ub -E4BU (U-box)	U	RING domain facilitates closed E2~Ub conformations and reveal the Arg residue from RING as a 'Linchpin' interacting to the close proximity of E2s 3 ₁₀ helix (Fig. 5c)	-	129
UbcH5a~Ub-RNF4	R2	Tyr193' of RING recognition to Ub Ile-36 patch	4AP4	130
UbcH5b~Ub-Birc7	R2	Phe296' and Arg294' of RING recognition to Ub Ile-36	4AUQ	131
UbcH5b~Ub-pCbl-b (Phosphorylated)	R	pTyr363 from non-RING element releases the auto-inhibition state and favors the Ub catalytic transfer	3ZNI	132, 145
III. Structures of E2~Ub/Ubl-E3-Substrate (RING) -snapshot of ubiquitin ligation				
UbcH5c - Bmi1/RING1B-Substrate Nucleosome	R2'	UbcH5c involved in substrate recognition, which presents the site-specific Lys119 of H2A to its acceptor residue.	4R8P	136
Ubc12~NEDD8 -Rbx1-substrate DCN1	R	NEDD8-type ubiquitination presents the ligation step and Trp35 residue contacts Ubl NEDD8 (Fig.6b)	4P5O	135
UbcH10~Ub-APC ^{CDH1} substrate (cryo-EM)	R	Two studies conclude that APC11's WHB domain interacts as a secondary E2 binding domain that trigger RING E3-mediated site-specific substrate ubiquitination.	EMD -2929 2925/6	137, 138
P-IV, Paper IV in the current thesis. Structures on HECT and RBR type -E3 ligases are omitted.				
εR, R2 and R2' represents for RING monomer, homodimer and heterodimer, respectively. U - U-box				

Non-covalent interactions between E2 helix2 surface and donor Ub

The interaction between donor Ub and the E2 conjugating enzyme was first reported biochemically, which emphasized the critical role of the hydrophobic Ile44 surface patch on the donor Ub contacting the E2 Helix2 surface¹²³. Additional findings using a combination of NMR CSP analysis and computational docking confirmed the requirement of essential non-covalent interaction between the donor Ub and E2 (ref¹²⁸), and furthermore provided the mechanistic details of how E2 (in this study – Ube2S) interaction with donor and acceptor Ub facilitates Lys11 linkage specific Ub chains¹²⁸. Solution-based studies on the E2~Ub thioester conjugate report a ‘closed’ conformation where Ub non-covalent interactions with the E2 helix2, provide a favorable position for Ub transfer, while ‘open’ conformations show no interactions between Ub and E2 except for the linkage site⁸ (**Fig. 5d**). Continued structural analysis in solution studies suggest that interactions between the donor Ub and RING domain are important for shifting the ensemble population of the E2~Ub conjugate towards a closed or restricted state¹²⁹.

(This space intentionally left blank.)

4.6. Structures of E2~Ub-E3 -Molecular basis of RING E3s stimulate Ub transfer

Recent ternary structure complexes of E2~Ub-RING E3 have shed more light on possible molecular mechanisms of RING-type promotion of Ub transfer, with UbcH5 E2~Ub bound to RNF4 (ref¹³⁰), Birc7 (ref¹³¹) or CBL-B¹³² RING domains (**Table 4, II**). These ternary structure complexes were made possible by carefully engineered mutations to stabilize E2~Ub conjugate to avoid the dissociation of the conjugate upon complex formation with RING E3 and to maintain the ternary complexes during the crystallization process. Such mutations include C85K of UbcH5a (ref¹³⁰), N77A from 'HPN' motif in UbcH5b (ref¹³¹), and the S22R mutation, which prevents Ub to bind at the E2 'backside'⁹². These ternary structural complex studies reveal common features of Ub recognition and activation by RING E3 ligases (**Table 4, II** and **Fig. 6b**). RING-type E3 ligases work either with monomer, homodimer or heterodimer RING domains (section 4.4). Examples of monomeric^{129,132} and homo-dimeric RING domains^{130,131} show that aromatic or charged residues from E3-RING interface are important in the recognition and in the positioning of donor Ub's Ile36 patch (**Fig. 6b** and **Fig. 5b**), and thereby the most favorable state for the catalytic Ub transfer¹²⁹. Identifying the specific molecular details in an individual ubiquitylation module would be important for understanding specificity and consideration of therapeutic strategies⁷⁹.

Regulation of RING E3 activity through auto-inhibition

From the aforementioned crystal structures of the ternary complexes it is becoming clear that the RING E3s not only create a scaffold to bind E2~Ub and substrate, they also mediate a network of recognition, and position Ub in a favorable state for Ub transfer. Several reports provide evidence that the regulation of RING E3 activity modulates the associated ubiquitination process (reviewed in ref¹³³). The different forms of regulation for E3 include auto-inhibition by other E3 domains than RING, as well as release of auto-inhibition through ligand-binding or phosphorylation^{132,134}. For example, the ligand iso-ADP ribose binds at the interface of the RING and WWE domains of RNF146, inducing a conformational change that releases the auto-inhibition and stabilizes the conformation mediating the RING binding to E2 (ref¹³⁴). A conformational change is also involved in the phosphorylation-based release of auto-inhibition, where the LHR, an element that is located adjacent to RING domain of c-CBL E3 ligase, undergoes phosphorylation in the conserved Tyr (**Fig. 6b**). This modification induces a conformational change in both the catalytic RING and in the substrate binding domain, and the phosphorylated residue further primes the donor Ub recognition, thereby activating the Ub transfer¹³². Such carefully regulated E3 activity likely prevents the formation of spurious substrate ubiquitination.

Multisite E2 binding is essential for RING E3 ligases to mediate site-specific substrate ubiquitination

The ternary complex structures of E2~Ub-E3 clearly manifest that the RING E3 ligases interaction promote catalytic transfer of Ub to substrate. Nonetheless, how individual RING E3 ligases select their substrate for Ub modification, and how RING E3 drives the substrate close to the vicinity of active site of E2 enzyme to accomplish site-specificity remain big questions to be addressed. Whilst they still cannot be fully answered, recently, three E2~Ub/Ubl-RING-substrate structures have provided important details on the Ub ligation step (**Table 4, III**). A first example is provided by NEDD8 type ubiquitylation of CUL1 substrate, where RBX1 monomeric RING ligase binds to Ubc12~NEDD8 and extensively interacts with donor Ub by a set of charged residues working as a 'Lever', and the aromatic residue RBX1-Trp35 acts as a 'pivot'¹³⁵ (**Fig. 6b**). Furthermore, co-E3 (co-activator of E3) protein DCN1 binds to the substrate CUL1 and to the acetylated E2, which then brings RBX1 E3 ligase to the close proximity for the catalytic ligation step¹³⁵. The second example of an E2-E3-substrate complex is the PRC1 ubiquitylation module, composed of the hetero-dimeric RING1B-Bmi1 in complex with UbcH5c and the substrate Nucleosome core, H2A¹³⁶. Here, the structure shows how the RING E3 ligase orients the E2 to facilitate ubiquitination. Remarkably, this is achieved by additional interaction from DNA to the c-terminal helix of E2 (**Fig. 5c**). The third example, the structure of the APC (anaphase-promoting complex) RING E3 ligase mediating cell-cycle regulator proteins ubiquitylation¹³⁷, presents details on ubiquitination of numerous lysine on a single substrate. With its 19-subunit APC complex, the subunits of monomeric APC11's RING in complex with UbcH10, and additionally APC2's WHB domain to UbcH10 backside interface, together with substrate this structure presents a snapshot of Ub ligation¹³⁷. This exciting study also presents the advantage of using the cryo-EM method and provides the understanding of a larger protein assembly in this transiently dynamic system. Studies on the same APC ubiquitylation system determined at 3.6 Å resolution using cryo-EM, further add knowledge about the binding interfaces in their ubiquitination complexes, including the phosphorylation based regulation of the ubiquitylation process¹³⁸.

All the aforementioned E2~Ub-E3-substrate complexes present interesting sides of the Ub ligation towards the substrate, emphasizing the importance of the E2s additional or backside interactions that are distant from RING domain interfaces, to achieve the site-specific ubiquitylation on the target substrate. With this collection of new structural insights, a molecular basis is emerging for the specific ubiquitination activity in RING E3 ligases.

4.7. Current clinical drug targets in ubiquitination system

Dysregulation in Ub and Ubl conjugation pathways and in the ubiquitin-proteasome system (UPS) lead to cellular defects that can result in pathological diseases like neurodegenerative disorders, viral diseases, autoimmunity and cancer (reviewed ¹⁴⁶).

Proteasome and E2-conjugating enzyme inhibition

The first compelling example of small molecule drugs altering the UPS and UBL conjugation pathway is the proteasome inhibitor bortezomib (Velcade), which is used for the treatment of multiple myeloma or mantle cell lymphoma^{147,148}. However, the efficacy of bortezomib decreases during the treatment and so the search for new targets and better drugs is still ongoing. A possible strategy could be targeting the specific ubiquitination pathways at the initial level of activation and conjugation may potentially be an ideal therapeutic candidate. One such recent example is the Nedd8-activating enzyme (E1 enzyme) inhibitor, MLN4924, that blocks the transfer of Nedd8 to the cullin-RING ligases (CRL - one of the largest families of E3 ligases) resulting in the inactive CRL ligases¹⁴⁹. The mechanism of this small molecule adduct widely received much attention towards drug design that specifically targets a unique UBL conjugation pathway^{149,150}. Another similar approach is used for inactivating the CRL E3 ligases, where the small molecule inhibitor CC0651 has as target towards its specific cognate E2 enzyme, Cdc34A¹⁵¹ and the mechanism of inhibition reveals that the inhibitor CC0651 binds in the close vicinity of E2 helix2 surface¹⁵², thereby forming an inhibitory complex that prevents the catalytic transfer of donor Ub to substrate.

Could E3-ligases be drug targets?

The interaction of the E3 ligases with their different substrates is very specific: this consideration might turn E3 ligases into interesting drug targets, not only because of their involvement in pathological disease, but also because targeting the specificity in their interactions – without affecting other E3 ligases - would potentially decrease the side effects⁷⁹. Therefore, molecular investigations of how individual E3 ligase establish substrate specificity may unravel important mechanisms and provide molecular insights into their pathophysiological roles, providing novel therapeutic paths for drug design that inhibit or enhance the individual ubiquitylation system.

4. 8. TRIM family of RING-type E3 ubiquitin-ligases

Multimodular TRipartite Motif (TRIM) proteins form the largest family of RING-type E3 ligases and play important roles in a variety of cellular process, including regulation of cell cycle progression, differentiation, development, oncogenesis, apoptosis, and innate immunity¹⁴⁶. TRIM5 is a well-characterized TRIM member shown to have dual activities in restricting viral infections and in innate immune signaling, following infection by retroviruses. The gene encoding human *TRIM5* belongs to a TRIM cluster on chromosome 11p15 along with its paralogs *TRIM21*, *TRIM22*, *TRIM34*, and *TRIM6* (ref ¹⁴⁶). Despite several important TRIM5 studies, there is a dearth of structural knowledge of the TRIM proteins regarding the mechanism of substrate ubiquitination as well the specificity of interactions with E2 conjugating enzymes. Recent structural studies of TRIM proteins present many individual domain structures, but the structural overview of how the TRIM multimodular system works together to mediate ubiquitination remains largely absent. The aim of the following sections is to briefly introduce the recent structures in TRIM proteins and demonstrate how they provide interesting clues towards their specific molecular mechanism and selectivity functions.

Structural features of the tripartite motif

The TRIM consists of several Zn^{2+} binding domains, namely RING and one or two B-box domains (1/2), and a Coiled-coil region (**Fig. 7a**), altogether also denoted RBCC family. At the C-terminus, various substrate-binding domains provide substrate specificity. The TRIM proteins are grouped based on the number of B-boxes and on the identity of the substrate binding domain, thus dividing TRIM proteins into nine subclasses^{153,154}. Many TRIM members fall in subclasses containing a C-terminal substrate binding PRY-SPRY domain¹⁵⁵, which is considered to recognize receptors involved in the detection of higher order oligomers, such as viral capsid proteins.

While the RING domain is ubiquitously present in E3s, TRIM proteins exclusively contain Zn^{2+} binding B-box domains termed B-box type 1 (B-box 1) and B-box type 2 (B-box 2). Different B-box types describe distinct patterns of conserved Cys and His residue arrangement in Zn^{2+} binding (**Fig. 7b**). Some TRIMs only contain a B-box 2 domain, such as TRIM5 α and its highly homologous TRIMs, including TRIM21. Several structures of B-box have revealed a $\beta\beta\alpha$ fold similar to that of the RING domain (**Fig. 7c**). While the RING domain binds to E2 UBC, the B-box has been shown to be essential for higher-order self-association, and thereby mediate substrate ubiquitination^{156,157}. However, the mechanistic details of how this is achieved remains to be investigated. A similar structural fold observed in RINGs and B-boxes suggests that these two functional domains are a product of evolution from a

common ancestral domain (**Fig. 7b, c**). The Coiled-coil (CC) region completes the tripartite motif; the structural organization of CC region comprises three α -helices, where the long helix in the middle possesses hydrophobic residues that most commonly Leu residues, mediating the dimerization within the TRIM itself (**Fig. 7c**), and is also essential for higher-order oligomer formation¹⁵⁷.

The C-terminal substrate-binding domain

The C-terminal domain conveys the specificity of TRIM members in interactions with its cognate substrates. Most of TRIM family members, such as TRIM5 and TRIM21, possess the PRY-SPRY domain, which assumes an immunoglobulin-like fold (**Fig. 7a**). Several other family members contain PHD-BROMO, TIF1, IGF, ARF or NHL, while some contain combinations of substrate-binding domains. Taken together, the presence of variable substrate-binding domains in TRIMs determines ubiquitination of a wide array of substrates.

4.9. TRIM21 -the cellular role

TRIM21, also referred to as Ro52, plays a pivotal role in the pathogenesis of autoimmunity. Autoantibodies against TRIM21 have been detected in different autoimmune diseases. More particularly, TRIM21 is a major auto-antigenic target in Sjögren's syndrome (SS) and systemic lupus erythematosus (SLE)¹⁵⁸. Functional studies of TRIM21 interaction with substrates (see **Table 5**) collectively indicate that it regulates several molecular pathways by its function as an E3 ligase in substrate ubiquitination. TRIM21 has been shown to ubiquitinate several IFN regulatory factors (IRFs) including IRF3, IRF5, IRF7 and IRF8 (ref ¹⁵⁹⁻¹⁶²), and thereby acting as a negative regulator of innate immune signalling.

Furthermore, TRIM21 regulates the activity of many transcription factors that are directly related to cytokine production and oncogenic or tumor suppressor proteins, a function, which could provide a molecular connection between autoimmune diseases, inflammation and cancer. Increased risk of lymphomas in SLE patients has been widely noted¹⁶³, and patients with SS have been found to have a higher prevalence of non-Hodgkin's lymphomas, possibly due to antigenic simulation in lymphocytes by autoantigens¹⁶⁴. Although experimental studies support the notion of TRIM21- substrate interaction through its PRY-SPRY domain, many molecular questions remain to be answered. Importantly, details on how site-specificity and specific Ub topology might be achieved by TRIM21 and how specific E2s interactions may govern specificity need to be delineated.

Table 5: Lists of substrates undergoing TRIM21-mediated ubiquitination

TRIM21-mediated substrate ubiquitination		Functional mechanism & Ub linkage type	Refs
Kinase inhibitor, p27		Upon phosphorylation of residue Thr187 of p27, the TRIM21 mediates Lys48-linked polyubiquitination on p27	165
Transcription factors	IRF3	Targets IRF3 for degradation, and phosphorylation of Tyr393 of TRIM21 regulates the interaction	166, 167
	IRF5	TRIM21-mediated proteasomal degradation of IRF5	168
	IRF7	TRIM21-mediated proteasomal degradation of IRF7	169
	IRF8	TRIM21-mediated IRF8 ubiquitination enhances transcription. In another study, p62 interaction regulates the TRIM21-IRF8 interaction and ubiquitination	159, 170
	C/EBP α	Lys48 linked poly-ubiquitination and degradation	171
Kinase activator, IKK β		Negative effect on NF- κ B activation via IKK β mono-ubiquitinate resulting in its subcellular translocation and subsequently autophagic degradation	172 173
DDX41		Lys48 linked polyubiquitin on Lys9 and Lys115 of DDX41, and thus proteasomal degradation	174
GMPS		Mono-ubiquitination on Lys182 of GMPS alters the subcellular localization of GMPS, and thereby regulates the p53 stability	175

4.10. TRIM interactions with subsets of E2s

Although the specificity line of E2-E3 pairs are addressed in few RING E3 ligases both at the molecular and functional level (**Table 4**), little is known about the molecular phenomena defining E2 specificity by TRIM E3 ligases. However, the first glimpses of E2-E3 pairs in TRIMs using yeast two-hybrid studies reveals that TRIM E3 ligases prefer to interact with UBE2D and UBE2E families of E2s¹²². The class III UBE2E family holds N-terminal extensions from the UBC core domain comprising three members UBE2E1, UBE2E2 and UBE2E3 (**Table 3** and section 4.3). The UBE2E1 interaction with TRIMs is widely noted and the E2 counterpart is strictly nuclear enzyme, as the Ub conjugated UBE2E1 complex interacts with Importin-11 for its translocation from the cytoplasm to the nucleus¹⁷⁶.

Multimer Hetero-interactions between TRIM21 and TRIM5 α -Regulation within the family

TRIM hetero-interactions are primarily observed among closely related TRIM proteins, further diversifying TRIM functions and regulation¹⁷⁷. It is evident that some

TRIMs that lack RING domain, such as TRIM44 and TRIM29, associate with closely related TRIMs containing RING catalytic domain, and thus mediate ubiquitination. TRIM hetero-interactions are also observed between two RING-containing TRIMs. Notably, TRIM21 and TRIM5 α interactions are among them. Both *in vivo* and *in vitro* experiments in agreement reveal that TRIM21 specifically interacts with UbcH5b, mediating monoubiquitination on TRIM5 α which causes its for subcellular translocation, and subsequent deubiquitinase by YopJ¹⁷⁸ that modulate its activity.

TRIM21 synthesizes unanchored Lys63-linked polyubiquitin chain

Several TRIMs, including TRIM5, TRIM25 and TRIM21, have been shown to synthesize unanchored Lys63-linked polyubiquitin chains. This initiates a signaling cascade that activates the transcription factor NF- κ B, thus generating an immune response against infection. In support of this, TRIM21 RING- domain interaction with the heterodimeric E2, Ubc13-MMS2, leads to unanchored Lys63-linked polyubiquitin chain synthesis *in vitro*¹⁷⁹, which in turn is a stimulus of NF- κ B signaling pathway.

TRIM21 undergoes auto-ubiquitination by interacting with subsets of E2

Auto-ubiquitination of some E3 ligases has been shown to occur using *in vitro* experiments. This includes N-terminal auto-ubiquitination of E3 by Ube2W, which acts like a pseudo-substrate recognized by specific E2 enzymes, or primes triggering and signaling for Lys63 polyubiquitin chains by the heterodimeric E2s. In the context of the N-terminal monoubiquitination, it is noteworthy to mention that Ube2W enzyme lacks the aspartic acid residue that plays a pivotal role for the incoming substrate lysine in the Ub ligation step (**Fig. 5c**). This explains why Ube2W prefers functioning in N-terminal Ub attachment instead of the lysine site-specific ubiquitination. The well characterized RING homodimer RNF4, as well as heterodimer BRCA1-BARD1 undergo mono-ubiquitination by the Ube2W enzyme that triggers the Lys63-linked polyubiquitination chains on RNF4 and BRCA1 by heterodimeric E2 Ubc13-MMS2 (ref ^{118,180}). Likewise, the RING domain from TRIM21 interacts with Ube2W for mono-ubiquitination and subsequently the heterodimeric E2 Ubc13-MMS2 interacts to form Lys63-linked polyubiquitin chains¹⁸¹. Furthermore, *in vitro* experiments suggest that the ubiquitinated-TRIM21 undergoes deubiquitinating by Poh1, a DUB enzyme¹⁸¹, in order to modulate the TRIM21 activity towards the viral infection neutralization direction¹⁸². It is also tempting to posit that the TRIM21 conjugated monoubiquitination by Ube2W, could be physiologically used to prevent spurious ubiquitination-mediated immune signaling. In context of E2 interactions, it is evident that the TRIM21 interacts with a subsets of E2 enzymes, however, selectivity at the molecular basis remains to be explored.

Figure 7. Structural features of TRIM family members with a focus on TRIM21. (a) Domain structure of TRIM proteins. Variability in domains between different family

members is shown (dotted line). TRIM21 domain structure is shown and boundaries are highlighted. **(b)** Primary amino-acid sequences of the RING and B-box 2 domains of TRIM21 with residues coordinating to Zn^{2+} binding, together with homology sequence logo for residues 1-130 of TRIM21. The relevant size of amino acid letters in each stack indicates the residues that are the most prevalent in that position. **(c)** Model of TRIM anti-parallel dimer of B-box-Coiled-coil-PRY-SPRY region, deriving from superimposition of crystal structures of B-box-Coiled-coil¹⁸⁶ and Coiled-coil-PRY-SPRY¹⁸⁷ (C_α atoms of Coiled-coil region from PDB codes- 4TN3, 4CG4). **(d)** Surface representation model of full-length TRIMs anti-parallel dimer, with a tentatively positioned TRIM model from (c), and E2-Ub-RING (**Paper IV**). The TRIM hexamer oligomer complex is based on the model obtained from cryoEM projection map^{157,183}. Previously identified distance measure of TRIM full-length of 170 Å is portrayed^{157,183}.

Multisite E2 interactions by TRIM E3 ligases

The regulation of the E2 enzymes by additional regions that are distant from the E1/E3 overlapping interface evidently forms a basis for functional diversity and control complexity (section 4.6). In the TRIM domain-containing E3 ligases interaction, how this additional regulation works is hitherto unknown, but clues into this process are provided by mutational studies on the UbcH5c-TRIM32 interface¹⁸⁴. The residues of the C-terminal helix of UbcH5c (depicted on **Fig. 5c**) are essential for interaction with TRIM32 and ensuing substrate ubiquitination, since truncation of these E2 residues abolishes TRIM32 activity¹⁸⁴. However, which part of the TRIM region interacts and what molecular mechanism mediates this additional interaction on E2 are two critical questions that are still unanswered.

TRIMs higher-order oligomer formations

Structural investigations of TRIM-mediated ubiquitination for TRIM5α and homologous TRIMs show that E3 ubiquitin-ligase activity requires higher-order self-association (**Fig. 7c, d**), based on three structural domain features: (i) the B-box 2 domain facilitates higher-order assembly^{156,157}, (ii) the Coiled-coil region facilitates the antiparallel dimerization of TRIMs^{183,185}, as shown in recent structures comprising the Coiled-coil region with B-box 2 (ref¹⁸⁶) and PRY-SPRY domains¹⁸⁷ (**Fig. 7c**), and (iii) the L2 linker from the C-terminal end of the Coiled-coil domain mediates positioning of the substrate-binding domain¹⁸⁸ in the vicinity of proximal RING domain¹⁸⁷, thereby providing the close distance for substrate ubiquitination.

Consistently with the hexagonal model of higher-order self-assemblies, the surface representation of the full-length TRIMs higher-order oligomer is depicted in **Fig. 7d**, with a tentatively positioned TRIM full-length and E2-Ub conjugate onto the

model obtained from cryoEM projection map^{157,183}. Based on the previously reported distance of TRIM full length of 170 Å, this model could be confidently created; however, more investigations of how this represents in an individual TRIM E3 ligases, would be of much interest. Further biochemical analysis of TRIM5α and homologous TRIMs, reveal that the RING domain is a prerequisite for efficient higher-order oligomer formations¹⁸⁹. The most recent structure of the RING domain from TRIM5α in complex with Ubc13 suggests that RING dimerization is essential for mediating Lys63-linked ubiquitination¹⁴³. To conclude, while the most recent structural findings have revealed the conserved structural signature of TRIM proteins forming anti-parallel dimerization mediated by the Coiled-coil region¹⁷⁷, the formation of higher-order oligomers in other TRIMs remains to be explored.

Chapter 5.

Methodology

To provide a mechanistic understanding of the macromolecular structure and complexes, and their associated conformational changes at the molecular level, several biophysical techniques were used in this study. The structural methods employed were selected and combined so as to address the particular complexity of the research project, including X-ray crystallography, Small-Angle X-ray Scattering (SAXS) and NMR spectroscopy. In this chapter, the different biophysical techniques that have been used in this thesis are briefly introduced.

5.1. Recombinant protein production in *E.coli*

In order to perform structural studies, it is necessary to obtain a large amount of proteins. Firstly, the gene of interest is cloned using Ligation Independent Cloning (LIC) into pET based LIC vectors with a T7 promoter region. These vectors are then subjected to recombinant protein expression in BL21 (DE3) *E.coli* expression strains. Protein purification is carried out in many different steps in order to obtain pure and homogenous monodisperse protein samples, including the use of affinity chromatography and size-exclusion chromatography.

Protein stability and Characterization

Before performing any structural study, it is fundamental to select the right buffer to maximize protein stability and to assess the sample purity. Several biophysical techniques are available to achieve these tasks, such as Circular Dichroism, Differential static light scattering and Analytical Gel Filtration. The last technique is also useful when studying protein-protein interactions, which can also be studied using Isothermal Titration Calorimetry.

Circular Dichroism (CD) spectroscopy is a technique used to investigate secondary structure content and thermal stability as a function of wavelength and temperature. The method is based on measuring the difference in absorption of left and right circularly polarized light by optically active chiral molecules in the protein sample. For secondary structure content determination, measurements are run through the far-UV range (190 – 240 nm), where the CD signals arise primarily from the peptide bonds of the polypeptide backbone. Moreover, thermal stability assessments and the buffer optimization screenings are performed using the ‘**Differential static light scattering**’ method, where protein stability is tested in hundreds of different buffer conditions and at different temperatures.

In this study, purified proteins samples from small-scale expression tests were immediately subjected to **Analytical Gel Filtration** (AGF). Determining the elution volume of the protein from AGF allows one to directly infer the oligomeric state, the molecular mass of the protein, and to assess the monodispersity of the sample in the selected buffer. In fact, this technique not only allows one to detect the formation of protein oligomers but also that of moderate/high-affinity protein complexes, when different interacting proteins are mixed together. Although AGF is a good tool to screen the presence of protein complexes, the measure of quantitative kinetic and thermodynamic parameters related to the binding, such as enthalphy, entropy, affinity and stoichiometry, are provided by other methods such as **Isothermal Titration Calorimetry** (ITC). ITC is a technique which is based on the measurement of energy released or absorbed upon a interaction between two molecular systems at a constant temperature.

5.2. Integrated structural studies

The structural knowledge of a proteins three-dimension (3D) -structural architecture and analysis provides a molecular basis for understanding its function and thus gives interesting clues about its mechanism of action. Currently, three methods provide the macromolecule structures at atomic resolutions, these are X-ray crystallography, NMR spectroscopy and single particle cryo-electron microscopy. SAXS also provides architectural information, abliet at a lower resolution for larger macromolecular entities, such as high molecular weight proteins and complexes. It is also important to realize that understanding the peculiarities associated with macromolecule complexes is not evident from a single biophysical technique, but from the carefull combination of structural methods which provide key insights from various perspectives. For instance, the atomic resolution structures determined from X-ray crystallography can be complemented with the low resolution information obtained from SAXS. This combination can thus be used to validate the structural shape in solution, and also possibly provide the conformational states of the structural regions that are too flexible to be captured by X-ray crystallography. Likewise, the sample characterization from NMR and SAXS provides the degree of a structural foldedness in the macromolecule, as well as information about the dispersity of the protein preparations.

The following parts briefly describe the above-mentioned structural techniques used in this study. **Fig. 8** schematically illustrates the technical steps in the experimental workflow, with examples provided from this study, while **Table 6** below briefly compares the techniques from a practical viewpoint.

Table 6. Comparison of structural techniques

Feature	X-ray Crystallography	NMR spectroscopy	SAXS
Used in	Paper I, II, IV	Paper I, II, III, IV	Paper II, IV
1. Provides	The most widely used method to determine 3D-macromolecular structures, including the complexes they form (Fig. 8a). The technique can provide atomic resolution information.	The dynamics of macromolecules, and conformational transitions as well as protein folding upon interactions with another protein (Fig. 8e), are its distinct advantages.	Size and shape of macromolecules and its complexes in solution. Capable of providing information of position and orientation of domains (Fig. 8d).
2. State	Crystalline state	In solution	In solution
3.Required	Mono-disperse and highly pure identical samples free from aggregation are required for all the structural techniques. However, in crystallography, the crystallization step often will remove impurities, and for NMR low levels of aggregate are not observed due to line broadening, while SAXS is the least tolerant to aggregation due to the dependence of the scattering on the square of the scattering particles molecular weight.		
4. Benefits of integrating biophysical methods	It may be necessary to use complementary methods (e.g. mutagenesis, NMR, SAXS) to determine if the protein conformation obtained from the crystalline state is impacted by crystallization.	Dynamic information at the atomic level could be mapped onto the crystal structure in order to understand the mechanistic features of the complexes.	Model building of the flexible regions that are not obtained from crystal structure (Fig. 8d) and ensemble modelling approaches.
5. Sample* conc.	Above 10 mg mL ⁻¹	8 mg mL ⁻¹	0.1 - 10 mg mL ⁻¹
6. Protein characterization Disordered & Flexibility	Disordered regions are typically inaccessible to crystallography, but the mobility of structure can be assessed from B-factors. Furthermore, the atom B-factor impacted from local and global disorder, and even occupancy of ligands.	NMR can very efficiently provide detailed insights by screening different buffers, and can structurally characterize the flexible region.	Buffer screening provides a means to assure monodisperse samples, and overall flexibility of the macromolecule can be assessed using Kratky plots.
7. Impact of the molecular size	Irrespective of molecular size, and as long as any flexible regions present do not interfere with the crystal lattice contacts, structure determination is possible once crystals are obtained.	Slow tumbling of large molecules leads to fast decay of NMR signal and line broadening and overlapping peaks in NMR spectra. Partial deuteration of large proteins facilitates measurements on larger proteins.	The scattering signal is proportional to the square of the volume of the scattering particle and only linearly with concentration; So large molecules can be studied at lower concentrations.

	X-ray Crystallography	NMR spectroscopy	SAXS
8. Feasibility of sample control	High-concentration of proteins are used to setup different crystallization conditions, although the observation of precipitation is informative, but at the expense of used material.	HSQC spectra are the ‘thumbprint’ of a protein that can be informative to ascertain right choice of buffer, as well as to check for folded and aggregated proteins.	R_g and $I(0)$ values and evaluation of the scattering in low ‘ q ’ ranges from measurements of concentration series provides immediate assessment of the sample quality.
9. Insights into Protein-protein complex	Provides high-resolution structures of the complex interface, provided that the binding constant is sufficient for the complex of interest to be stable enough to crystallize.	Per-residue analysis of the interface for both transient and weak-interacting complexes.	Information on the degree of association or oligomerization for dynamic complexes. Where data are available on the individual components, conformational changes upon complex formation might be obtained.
10. Protein structural heterogeneity	Heterogeneous systems are often masked or invisible within the confines of a crystal lattice.	Multiple peaks for the same nucleus is an indication, and NMR relaxation experiments provide information on conformational dynamics on multiple timescales with per-residue.	Time-resolved SAXS can provide information on the global conformation in the presence of multiple states.
11. Concerns	High concentration of proteins samples used to obtain the crystal structure, may alter conformations by non-biological interactions; furthermore, sparsely populated states are generally invisible in X-ray crystallography.	Care is required in the region of overlapping resonance in NMR spectra.	Over-interpretation of the inherently 1D data provided by SAXS in terms of 3D structures is possible. Stringent sample quality control, careful application of 3D modelling approaches and validation using complementary data is important.

	X-ray Crystallography	NMR spectroscopy	SAXS
12. Resolution	Is a measure of the level of detail present in the diffraction pattern (Fig. 8c), which is determined by the size and degree of perfection of the crystal. The highest angle data measured determines smallest distances between scattering centers that can be resolved. Higher resolution structures reveal more details (Fig. 8c), however the quality of the electron density map vary across the model since some parts are more/less ordered or have unique/several conformations.	In NMR, resolution refers to the ability to resolve individual resonance lines and depends upon the peak line-shape, commonly described simply in terms of the half-height line width.	SAXS data, like crystal diffraction data, are reciprocal space measurement and while SAXS data can be measured to quite high angles, the information content is limited by the rotational averaging due to the random orientations of the molecules in solution. In SAXS measurement, resolution generally refers to minimum measurement angle, which determines the longest distances that can be reliably determined.
<i>In vitro</i> to <i>In vivo</i> relevance	The obtained structural information at the molecular level could be useful to design drug targets, as the crystal structure typically shows one conformation, it is usually preferred by the pharmaceutical companies.	NMR structural determinations could be performed at solution conditions where physiologically relevant parameters (such as pH, ionic strength, presence of ligands) can be readily controlled. There is a size limit to the systems that can be studied due to slow tumbling rates and the resultant line broadening.	Similar to NMR, the measurements are made in solution where physiologically relevant parameters can be controlled. The molecular weight range that can be studied however is similar to crystallography; from the smallest of peptides to something as large as the influenza virus constructs can be feasibly studied.
To address the increasing biological complexity it is necessary to use several biophysical techniques that provide unique and complementary insights into understanding structure-function relationship; however, the challenge of integrative model complex systems should also require careful analysis and attention to the inherent experimental errors and limitations of different kinds of data to avoid mis- and over-interpretation.			

* Sample concentration used in this study

Figure 8. Experimental workflow through X-ray crystallography, SAXS and NMR spectroscopy. (a) An overview of X-ray crystallography approach to determine 3D protein structures, presented here with examples from this study. The steps involved are: (i)

crystallization of the protein of interest, (ii) X-ray diffraction and data collection, (iii) From the diffraction pattern to the electron density, and (iv) Structural model and analysis. Obtaining protein crystals is the foremost step, where the protein molecules must be arranged in an ordered three-dimensional array of molecules, leading into a periodic lattice. The smallest repeating units of a crystal represent the unit cell, arranged to a space group symmetry. The general procedure used here to crystallize a protein is the vapor diffusion method. Two crucial sequential steps occur to produce protein crystals: spontaneous nucleation and nucleated crystal growth. In this study, a carefully selected combination of seeding procedure to facilitate the crystal growth was followed as described in (b) insets. The diffraction resolution is shown for these crystals. Upon collection of the complete dataset based on the space group information, the next step is to obtain electron density map from the collected diffraction data, where the data contains the amplitude of diffraction spots but the information on the relative phase of the diffracted radiation is lost: this drawback is referred as “the phase problem”. In this study the phase problem is solved by method of choice, such as molecular replacement (Paper- I, II) and multiple anomalous dispersion (Paper- IV). Based on the diffraction resolution, the details of the electron density maps are provided, as previously described in Table 6 and shown in (c). For example, the density map for Tyr residue of UBE2E1 is defined in both cases at 1.4Å and 2.86Å resolution, while the 1.4Å resolution provides well-defined map, shown the model site with the 2Fo-Fc map (grey mesh) contoured at 1.5 σ . The final step of X-ray crystallography is to build the structural model with several rounds of refinement, such as examining the torsion angles of amino-acids by the help of Ramachandran plot. The quality of refinement is measured by the R-factor.

(d) An overview of the small-angle X-ray scattering approach to obtain the size and shape of macromolecules in solution. The scattering data from a macromolecule represent a time and ensemble average of randomly oriented structures, and that the rotational-averaging of 3D structures presents a 1D profile. The experimental scattering curve show the intensity of scattering by Log $I(q)$ versus q plot, (where $q = 4\pi \sin\theta/\lambda$). For a mono-disperse, particles of the same size and overall shape the Log-log plots generally show the characteristic flattening at $q_{\min}$ region and Guinier plots should be linear for $qR_g < 1.3$ as indicator for a good quality sample that is free from inter-particle interference and aggregation. Fourier transformation of the scattering profile to $P(r)$, the inter-atomic distances distribution function, provides the information on the shape, and together with the forward or zero angle scattering intensity, $I(0)$, that can provide the size of the macromolecule with high accuracy. From the example provided here, shows the reconstruction of UBE2E1 full-length dummy atom sphere models (*ab initio* shape envelopes) calculated using the $P(r)$'s as input. While the crystal structure of UBE2E1 full-length (PDB code - 3BZH) superimposes well with *ab initio* shape envelopes produced from

SAXS data (DAMMIF program), however, the flexible N-terminal extensions (residues 1-40) are visible in SAXS structural analysis.

(e) An overview of the NMR spectroscopy techniques used in this work, with the goal of obtaining the molecular basis of protein-protein interactions in solution. TROSY experiments of Isotope labeled ^{15}N protein samples are measured, to detect changes in the chemical environment for amino acids upon complex formation by adding increasing concentrations of the unlabeled interaction partner. TROSY experiments are two-dimensional spectra where each peak corresponds to a signal from an amide group, either from the backbone or from the side chains of the amino acids. The changes in the peak resonance position (peak shift) or intensity of the backbone amides (peak broadened) upon complex formation are graphically shown in insets here. In case of peak shifts, chemical shift perturbations (CSPs) could be calculated and mapped on the structure, defining the residues that are directly or indirectly affected by the binding. Here presented the spectra of ^{15}N UBE2E1 and titration with unlabeled TRIM21₁₋₉₁ construct.

Summary of Papers and Prospective

The Summary of Papers and Prospective associated with this thesis have been removed for copyright reasons.

Acknowledgments

Life is amazing with scientific research! I have enthusiastically enjoyed every moment of my doctoral studies. I wish to acknowledge the many wonderful people with whom I have interacted with and worked as a team, in particular, my sincere thanks to:

My supervisor, **Maria Sunnerhagen**, for providing me the best in every way at every stage throughout my PhD studies. I thank you deeply for sharing your knowledge and training me in this exciting research, and especially for the opportunities to work in the best-integrated structural biology. I am sincerely grateful to you for everything.

My co-supervisor, **Martin Moche**, It has been a great privilege to work with you and be trained. I admire your energetic and enthusiastic way of looking at things, special appreciation for providing me a great source of energy and raising my confidence.

My co-supervisor, **Marie Wahren-Herlenius**, I am grateful to you for all the discussion opportunities of our collaborative teamwork and constant encouragements.

Björn Wallner, Thank you so much for all the wonderful discussions and the collaborative teamwork. Special thanks for prodding me to look at things differently.

Patrik Lundström, I am grateful to you, especially for your excellent guidance with NMR-related experiments, and for pedagogically teaching.

Jill Trehwella, thank you for all the support you provided me, and special thanks for sharing your enthusiasm for research, and also for SAXS knowledge.

Don Parkin, you are an amazing person and a born-teacher. Thank you for everything and especially for training in SAXS instrument and sharing your experiences.

Christian Löw, You brought me to the structural biology, which I will acknowledge throughout my career and my heartfelt thanks for all inspiration and motivations.

Meri Montecchio, the most down-to-earth person I have ever met. Your astonishing attention to detail and perfection is something that I always admire. While I continue thinking for my last sentence, I express my sincere and warm thanks.

Cecilia Andresen, hard to acknowledge the loveliest person on this planet in few words! Many thanks for all the much-needed advice and help throughout my PhD.

Nikolaos Kyriakidis, I am amazed with your excellence and perfection! I am also glad that we worked mutually with a similar focus. Huge thanks for all the beautiful work.

Susanne Andersson, for being helpful with many things and for your timely inputs.

Louise Gustafsson, I am sincerely grateful to you for all your support.

Special thanks to some of my course teachers: *Alwyn Jones, Terese Bergfors, Per Hammarström, Mathieu Linares, Martin Magnuson, Per Jensen, Eva Bergqvist* and EMBO courses. I also thank *Uwe Muller* for his amazing BESSY BL14-1 and -2 beamlines, and *Adam Round* for excellent SAXS beamline at ESRF. Special thanks to my PhD mentor, *Jordi Altimiras*.

Sunnerhagen group, Amelie Wallenhammar, I deeply appreciate for all your support and kindness, wishing you all the best. *Emily Chamberlain*, dedicated and a clever person, thank you for everything and especially for your proofreading skills, and for organizing things for my dissertation. *Marie Roth*, for taking good care of many things, and *Tove Canerstam* wishing you all best on your Thesis work.

Former colleagues: Sara Helander, for all your help and energetic way of addressing things. *Malin Elven*, those beautiful days, thanks for all fun! *Tommy Larsson, Rebecca Anderson, Alexander Lepak, Emmy Österberg, Helena Malmrot, Sandra Weimberger, Javed M.E. Ziauddin, Vishnu Revuri, Jonna Luoma* – Thank you all for the teamwork! *Janosch Hennig*, for your discussions.

Protein Science Facility at KI: Helena Berglund and Tomas Nyman, Special thanks for always welcoming me and providing a learning atmosphere, and also for the extra bonus facilities. I also thank: *Eleonora Maccacaro, Ida Johansson, Emma Hansson*. Special thanks to the old SGC team, especially to *Marina Siponen*.

Wallner group, Claudio Mirabello, Thank you for all your professional and personal help, and for always listening to me. *Robert Pilstål*, Thanks for your contributions in our joint team-effort. Thanks to *Malin Larsson, Sankar Basu* and *Mika Gustafsson*.

Wahren group, I am thankful to *Alexander Espinosa, Aurelie Ambrosi, Maria Sjöstrand, Vilija oke, Sabrina Görgen* and *Sabrina Goergen*.

Lundstrom group, Alexandra Ahlner, It is always great to talk with you and thanks for all the supportive discussions. *Markus Niklasson*, thanks for all the nice company and especially at Tallberg, *Ivana Magic*, you are an amazing person, thanks for always being very caring. Thanks to *Sofie Cassman* and *Patricia Gullberg*.

Chemistry division at IFM, LiU: Leif B.G. Johansson (Lefte), many thanks for always making my days beautiful and for all fun chitchats. *Karin Magnusson*, it has been great talking with you, especially our Friday chats. *Annica Blissing*, special thanks to you. *Bengt-Harald Jonsson (Nalle)*, for contributing to a nice ambience at our division. *Magdalena Svensson* -I appreciate for your inputs. *Per Hammarström*, thanks for helping me with postdoc hunting. *Lars-Göran Mårtensson (Lasse)*, you're such an inspiring teacher. Many thanks for all help! *Uno Carlsson*, thanks for sharing your knowledge. I extend my humble thanks to *Sofie Nyström, Maria Lundqvist, Ann-Christu Brorsson, Anna Zimdahl Kahlin, Christine "Bisse" Dyrhagen, Lotta Tollstoy Tegler, Patrik Nygren, Maria Jonsson, Alexander Sandberg, Maria Thörnqvist, Martin Karlsson, Jutta Speda, Liza Bergkvist, Uwe Pielles, Mikaela Johansson, Peter Nilsson, Mathias Elgland, Marcus Bäck, Peter Kondrasson*,

Johan Dahlén and **Stefan Klinström**. I also thank our former colleagues: **Therese Klingstedt**, **Linda Helmfors**, **Veronica Sandgren**, **Daniel Kanmert**, **Rozalyn Simon**, **Daniel Sjölander** and **Ida Hafstrand**. Special thanks to **Gunnar Bergström**, **Robert Gustavsson** and **Jonas Christoffersson**.

I am grateful to **Patricia Wennerstrand** and **Robin Karlsson**, I must especially acknowledge for all your comforting talks.

MBB department at KI: Rosaria Gandini, I have always admired your dedication towards research and thanks for all the advice. Special thanks to **Fatma Guettou**, I always enjoy your company, especially the synchrotron time. **Pär N**, **Per M**, **Esben Q**, **Sing ML (Rachel)**, **Rozbeh J**, **Daniel MM**, **Rebecca D**, **Madeleine A**, **Mikaela R**, **Tobias K**, **Agnes Rinaldo-Matthis**, **Madhuranayaki T**, **Herwig S**, **Ann-Gred T**, **Filipa P** and other MBB colleagues- many thanks to all of you for your inputs.

Abeni Wickam, I am very thankful to your proofreading skills. Thanks to **Ranjith R**, **Arnaud L. F.**, **Annop E.**, **Andreas S.** and other IFM friends.

Sunnerhagen family, **Maria**, **Bo**, **Örjan** and **Hugo**, I deeply acknowledge all of you for giving me such a nice welcome whenever I visit your family. The time with your family in Skirö was beautiful; the late evening sunsets are still imprinted on my retina.

Moche family, **Martin**, **Carina**, **Carl-Axel** and **Elsa**, that beautiful home is still in my memory, millions of thanks to all of you for many great and fun conversations. **Martin**, I learned so many wonderful things from you and my sincere thanks for that.

Löw family, **Christian**, **Mirjam**, **Svea** and **Filip**, I am extraordinarily grateful to all of you and I look forward to meet you sometime in the future.

Phopase family, **Jaywant**, **Nalini** and **Shandilya**, thanks to all your well wishes, and it has been great talking with you all these years. My best wishes to Shandilya's future.

Devarajulu, **Ramakrishnan**, **Balaji**, **Krishnamurthy families**- huge thanks to all the great support that you all provided throughout these years! My special thanks to **Ram**.

My younger sister, **Mohana Lakshmi**, I could start the acknowledgement by calling you my family doctor. I know, we have rarely been formal, but all I want to say is you are a priceless asset to me and to our parents. Let us jointly make our parents proud!

My Parents, **Ananda padamanaban & Renuga Padamanaban**, my treasure and secret of living a happy life. From the bottom of my heart, I am extraordinarily grateful to you both for everything, and in particular, for teaching me the meaning and principles of life. The rest of this note, I will keep personal :)

REFERENCES

1. CRICK, F. Central Dogma of Molecular Biology. *Nature* **227**, 561–563 (1970).
2. Vickery, H. B. The origin of the word protein. *The Yale journal of biology and medicine* **22**, 387–393 (1950).
3. Wright, P. E. & Dyson, H. J. Linking folding and binding. *Curr. Opin. Struct. Biol.* **19**, 31–38 (2009).
4. Fischer, E. Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der deutschen chemischen Gesellschaft* **27**, 2985–2993 (1894).
5. Koshland, D. E. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proceedings of the National Academy of Sciences* **44**, 98–104 (1958).
6. Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).
7. Nussinov, R., Ma, B. & Tsai, C.-J. Multiple conformational selection and induced fit events take place in allosteric propagation. *Biophys. Chem.* **186C**, 22–30 (2014).
8. Pruneda, J. N., Stoll, K. E., Bolton, L. J., Brzovic, P. S. & Klevit, R. E. Ubiquitin in motion: structural studies of the ubiquitin-conjugating enzyme~ubiquitin conjugate. *Biochemistry* **50**, 1624–1633 (2011).
9. Thomas, M. C. & Chiang, C.-M. The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41**, 105–178 (2006).
10. Hahn, S. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* **11**, 394–403 (2004).
11. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nature reviews. Molecular cell biology* **16**, 129–143 (2015).
12. Plaschka, C. *et al.* Architecture of the RNA polymerase II-Mediator core initiation complex. *Nature* **518**, 376–380 (2015).
13. Kandiah, E., Trowitzsch, S., Gupta, K., Haffke, M. & Berger, I. More pieces to the puzzle: recent structural insights into class II transcription initiation. *Curr. Opin. Struct. Biol.* **24**, 91–97 (2014).
14. Bieniossek, C. *et al.* The architecture of human general transcription factor TFIID core complex. *Nature* **493**, 699–702 (2013).
15. Hernandez, N. TBP, a universal eukaryotic transcription factor? *Genes & development* **7**, 1291–1308 (1993).
16. Elmlund, H. *et al.* Cryo-EM reveals promoter DNA binding and conformational flexibility of the general transcription factor TFIID. *Structure* **17**, 1442–1452 (2009).
17. Nikolov, D. B. *et al.* Crystal structure of TFIID TATA-box binding protein. *Nature* **360**, 40–46 (1992).
18. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. Crystal structure of a yeast TBP/TATA-box complex. *Nature* **365**, 512–520 (1993).
19. Chitikila, C., Huisinga, K. L., Irvin, J. D., Basehoar, A. D. & Pugh, B. F. Interplay of TBP inhibitors in global transcriptional control. *Molecular Cell* **10**, 871–882 (2002).
20. Chasman, D. I., Flaherty, K. M., Sharp, P. A. & Kornberg, R. D. Crystal structure of yeast TATA-binding protein and model for interaction with DNA. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8174–8178 (1993).
21. Kim, J. L., Nikolov, D. B. & Burley, S. K. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365**, 520–527 (1993).
22. Tan, S., Hunziker, Y., Sargent, D. F. & Richmond, T. J. Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* **381**, 127–151 (1996).
23. Bleichenbacher, M., Tan, S. & Richmond, T. J. Novel interactions between the components of human and yeast TFIIA/TBP/DNA complexes. *J. Mol. Biol.* **332**, 783–793 (2003).
24. Liu, D. *et al.* Solution structure of a TBP-TAF(II)230 complex: protein mimicry of the minor groove surface of the TATA box unwound by TBP. *Cell* **94**, 573–583 (1998).
25. Anandapadamanaban, M. *et al.* High-resolution structure of TBP with TAF1 reveals anchoring patterns in transcriptional regulation. *Nat. Struct. Mol. Biol.* **20**, 1008–1014 (2013).
26. Nikolov, D. B. *et al.* Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**, 119–128 (1995).
27. Wollmann, P. *et al.* Structure and mechanism of the Swi2/Snf2 remodeller Mot1 in complex with its substrate TBP. *Nature* **475**, 403–407 (2011).

28. Kamada, K. *et al.* Crystal structure of negative cofactor 2 recognizing the TBP-DNA transcription complex. *Cell* **106**, 71–81 (2001).
29. Juo, Z. S., Kassavetis, G. A., Wang, J., Geiduschek, E. P. & Sigler, P. B. Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature* **422**, 534–539 (2003).
30. Butryn, A. *et al.* Structural basis for recognition and remodeling of the TBP:DNA:NC2 complex by Mot1. *Elife* **4**, e07432 (2015).
31. Papai, G. *et al.* TFIIA and the transactivator Rap1 cooperate to commit TFIID for transcription initiation. *Nature* **465**, 956–960 (2010).
32. Liu, W.-L. *et al.* Structures of three distinct activator-TFIID complexes. *Genes & development* **23**, 1510–1521 (2009).
33. Kou, H., Irvin, J. D., Huisinga, K. L., Mitra, M. & Pugh, B. F. Structural and functional analysis of mutations along the crystallographic dimer interface of the yeast TATA binding protein. *Molecular and cellular biology* **23**, 3186–3201 (2003).
34. Yatherajam, G., Zhang, L., Kraemer, S. M. & Stargell, L. A. Protein-protein interaction map for yeast TFIID. *Nucleic Acids Research* **31**, 1252–1260 (2003).
35. Nishikawa, J., Kokubo, T., Horikoshi, M., Roeder, R. G. & Nakatani, Y. Drosophila TAF(II)230 and the transcriptional activator VP16 bind competitively to the TATA box-binding domain of the TATA box-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 85–90 (1997).
36. Martel, L. S., Brown, H. J. & Berk, A. J. Evidence that TAF-TATA box-binding protein interactions are required for activated transcription in mammalian cells. *Molecular and cellular biology* **22**, 2788–2798 (2002).
37. Pham, A. D. & Sauer, F. Ubiquitin-activating/conjugating activity of TAFII250, a mediator of activation of gene expression in Drosophila. *Science* **289**, 2357–2360 (2000).
38. Mal, T. K. *et al.* Structural and functional characterization on the interaction of yeast TFIID subunit TAF1 with TATA-binding protein. *J. Mol. Biol.* **339**, 681–693 (2004).
39. Kotani, T. *et al.* A role of transcriptional activators as antirepressors for the autoinhibitory activity of TATA box binding of transcription factor IID. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7178–7183 (2000).
40. Bagby, S. *et al.* TFIIA-TAF regulatory interplay: NMR evidence for overlapping binding sites on TBP. *FEBS letters* **468**, 149–154 (2000).
41. Kokubo, T., Swanson, M. J., Nishikawa, J. I., Hinnebusch, A. G. & Nakatani, Y. The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. *Molecular and cellular biology* **18**, 1003–1012 (1998).
42. Goppelt, A., Stelzer, G., Lottspeich, F. & Meisterernst, M. A mechanism for repression of class II gene transcription through specific binding of NC2 to TBP-promoter complexes via heterodimeric histone fold domains. *EMBO J.* **15**, 3105–3116 (1996).
43. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
44. Hateboer, G. *et al.* TATA-binding protein and the retinoblastoma gene product bind to overlapping epitopes on c-Myc and adenovirus E1A protein. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8489–8493 (1993).
45. McEwan, I. J., Dahlgren-Wright, K., Ford, J. & Wright, A. P. Functional interaction of the c-Myc transactivation domain with the TATA binding protein: evidence for an induced fit model of transactivation domain folding. *Biochemistry* **35**, 9584–9593 (1996).
46. Fladvad, M. *et al.* N and C-terminal sub-regions in the c-Myc transactivation region and their joint role in creating versatility in folding and binding. *J. Mol. Biol.* **346**, 175–189 (2005).
47. Grove, A. MarR family transcription factors. *Curr. Biol.* **23**, R142–3 (2013).
48. Perera, I. C. & Grove, A. Molecular mechanisms of ligand-mediated attenuation of DNA binding by MarR family transcriptional regulators. *J Mol Cell Biol* **2**, 243–254 (2010).
49. Alekshun, M. N., Levy, S. B., Mealy, T. R., Seaton, B. A. & Head, J. F. The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 Å resolution. *Nature structural biology* **8**, 710–714 (2001).
50. Hong, M., Fuangthong, M., Helmann, J. D. & Brennan, R. G. Structure of an OhrR-ohrA Operator Complex Reveals the DNA Binding Mechanism of the MarR Family. *Molecular Cell* **20**, 131–141 (2005).
51. Dolan, K. T., Duguid, E. M. & He, C. Crystal Structures of SlyA Protein, a Master Virulence Regulator of Salmonella, in Free and DNA-bound States. *J. Biol. Chem.* **286**, 22178–22185

- (2011).
52. Kumarevel, T., Tanaka, T., Umehara, T. & Yokoyama, S. ST1710-DNA complex crystal structure reveals the DNA binding mechanism of the MarR family of regulators. *Nucleic Acids Research* **37**, 4723–4735 (2009).
 53. Birukou, I., Seo, S. M., Schindler, B. D., Kaatz, G. W. & Brennan, R. G. Structural mechanism of transcription regulation of the *Staphylococcus aureus* multidrug efflux operon *mepRA* by the MarR family repressor MepR. *Nucleic Acids Research* **42**, 2774–2788 (2014).
 54. Andréßen, C. *et al.* Critical biophysical properties in the *Pseudomonas aeruginosa* efflux gene regulator MexR are targeted by mutations conferring multidrug resistance. *Protein Science* **19**, 680–692 (2010).
 55. Lim, D. Crystal Structure of the MexR Repressor of the *mexRAB-oprM* Multidrug Efflux Operon of *Pseudomonas aeruginosa*. *Journal of Biological Chemistry* **277**, 29253–29259 (2002).
 56. Birukou, I. *et al.* The Molecular Mechanisms of Allosteric Mutations Impairing MepR Repressor Function in Multidrug-Resistant Strains of *Staphylococcus aureus*. *mBio* **4**, e00528–13–e00528–13 (2013).
 57. Kumaraswami, M., Schuman, J. T., Seo, S. M., Kaatz, G. W. & Brennan, R. G. Structural and biochemical characterization of MepR, a multidrug binding transcription regulator of the *Staphylococcus aureus* multidrug efflux pump MepA. *Nucleic Acids Research* **37**, 1211–1224 (2009).
 58. Bordelon, T., Wilkinson, S. P., Grove, A. & Newcomer, M. E. The crystal structure of the transcriptional regulator HucR from *Deinococcus radiodurans* reveals a repressor preconfigured for DNA binding. *J. Mol. Biol.* **360**, 168–177 (2006).
 59. Davis, J. R., Brown, B. L., Page, R. & Sello, J. K. Study of PcaV from *Streptomyces coelicolor* yields new insights into ligand-responsive MarR family transcription factors. *Nucleic Acids Research* **41**, 3888–3900 (2013).
 60. Perera, I. C., Lee, Y.-H., Wilkinson, S. P. & Grove, A. Mechanism for attenuation of DNA binding by MarR family transcriptional regulators by small molecule ligands. *J. Mol. Biol.* **390**, 1019–1029 (2009).
 61. Chang, Y.-M., Chen, C. K.-M., Ko, T.-P., Chang-Chien, M. W. & Wang, A. H.-J. Structural analysis of the antibiotic-recognition mechanism of MarR proteins. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1138–1149 (2013).
 62. Newberry, K. J., Fuangthong, M., Panmanee, W., Mongkolsuk, S. & Brennan, R. G. Structural Mechanism of Organic Hydroperoxide Induction of the Transcription Regulator OhrR. *Molecular Cell* **28**, 652–664 (2007).
 63. Sun, F. *et al.* Protein cysteine phosphorylation of SarA/MgrA family transcriptional regulators mediates bacterial virulence and antibiotic resistance. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15461–15466 (2012).
 64. Wilke, M. S. *et al.* The crystal structure of MexR from *Pseudomonas aeruginosa* in complex with its antirepressor ArmR. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 14832–14837 (2008).
 65. Chen, H. *et al.* scientific report. *EMBO reports* **11**, 685–690 (2010).
 66. Saridakis, V., Shahinas, D., Xu, X. & Christendat, D. Structural insight on the mechanism of regulation of the MarR family of proteins: high-resolution crystal structure of a transcriptional repressor from *Methanobacterium thermoautotrophicum*. *J. Mol. Biol.* **377**, 655–667 (2008).
 67. Kerr, K. G. & Snelling, A. M. *Pseudomonas aeruginosa*: a formidable and ever-present adversary. *J. Hosp. Infect.* **73**, 338–344 (2009).
 68. Poole, K. Multidrug efflux pumps and antimicrobial resistance in *Pseudomonas aeruginosa* and related organisms. *J. Mol. Microbiol. Biotechnol.* **3**, 255–264 (2001).
 69. Mima, T., Sekiya, H., Mizushima, T., Kuroda, T. & Tsuchiya, T. Gene cloning and properties of the RND-type multidrug efflux pumps MexPQ-OpmE and MexMN-OprM from *Pseudomonas aeruginosa*. *Microbiol. Immunol.* **49**, 999–1002 (2005).
 70. Blair, J. M. A. & Piddock, L. J. V. Structure, function and inhibition of RND efflux pumps in Gram-negative bacteria: an update. *Curr. Opin. Microbiol.* **12**, 512–519 (2009).
 71. Poole, K. & Srikumar, R. Multidrug efflux in *Pseudomonas aeruginosa*: components, mechanisms and clinical significance. *Curr Top Med Chem* **1**, 59–71 (2001).
 72. Poole, K. *et al.* Expression of the multidrug resistance operon *mexA-mexB-oprM* in *Pseudomonas aeruginosa*: *mexR* encodes a regulator of operon expression. *Antimicrob. Agents Chemother.* **40**, 2021–2028 (1996).

73. Adewoye, L., Sutherland, A., Srikumar, R. & Poole, K. The mexR repressor of the mexAB-oprM multidrug efflux operon in *Pseudomonas aeruginosa*: characterization of mutations compromising activity. *J. Bacteriol.* **184**, 4308–4312 (2002).
74. Alekshun, M. N., Kim, Y. S. & Levy, S. B. Mutational analysis of MarR, the negative regulator of marRAB expression in *Escherichia coli*, suggests the presence of two regions required for DNA binding. *Mol. Microbiol.* **35**, 1394–1404 (2000).
75. Duval, V., McMurry, L. M., Foster, K., Head, J. F. & Levy, S. B. Mutational analysis of the multiple-antibiotic resistance regulator MarR reveals a ligand binding pocket at the interface between the dimerization and DNA binding domains. *J. Bacteriol.* **195**, 3341–3351 (2013).
76. Chen, H. *et al.* The *Pseudomonas aeruginosa* multidrug efflux regulator MexR uses an oxidation-sensing mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13586–13591 (2008).
77. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479 (1998).
78. Pickart, C. M. & Eddins, M. J. Ubiquitin: structures, functions, mechanisms. *Biochim. Biophys. Acta* **1695**, 55–72 (2004).
79. Lipkowitz, S. & Weissman, A. M. RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. *Nature reviews. Cancer* **11**, 629–643 (2011).
80. Hoeller, D., Hecker, C.-M. & Dikic, I. Ubiquitin and ubiquitin-like proteins in cancer pathogenesis. *Nature reviews. Cancer* **6**, 776–788 (2006).
81. Wilkinson, K. D. Ubiquitin: a Nobel protein. *Cell* **119**, 741–745 (2004).
82. Li, W. *et al.* Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. *PLoS ONE* **3**, e1487 (2008).
83. Metzger, M. B., Pruneda, J. N., Klevit, R. E. & Weissman, A. M. RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochim. Biophys. Acta* **1843**, 47–60 (2014).
84. Berndsen, C. E. & Wolberger, C. New insights into ubiquitin E3 ligase mechanism. *Nat. Struct. Mol. Biol.* **21**, 301–307 (2014).
85. Deshaies, R. J. & Joazeiro, C. A. P. RING domain E3 ubiquitin ligases. *Annu. Rev. Biochem.* **78**, 399–434 (2009).
86. Zheng, N., Wang, P., Jeffrey, P. D. & Pavletich, N. P. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* **102**, 533–539 (2000).
87. Huang, D. T. *et al.* Structural basis for recruitment of Ubc12 by an E2 binding domain in NEDD8's E1. *Molecular Cell* **17**, 341–350 (2005).
88. Huang, L. *et al.* Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. *Science* **286**, 1321–1326 (1999).
89. Eletr, Z. M., Huang, D. T., Duda, D. M., Schulman, B. A. & Kuhlman, B. E2 conjugating enzymes must disengage from their E1 enzymes before E3-dependent ubiquitin and ubiquitin-like transfer. *Nat. Struct. Mol. Biol.* **12**, 933–934 (2005).
90. Dikic, I., Wakatsuki, S. & Walters, K. J. Ubiquitin-binding domains - from structures to functions. *Nature reviews. Molecular cell biology* **10**, 659–671 (2009).
91. Bode, A. M. & Dong, Z. Post-translational modification of p53 in tumorigenesis. *Nature reviews. Cancer* **4**, 793–805 (2004).
92. Brzovic, P. S., Lissounov, A., Christensen, D. E., Hoyt, D. W. & Klevit, R. E. A UbcH5/ubiquitin noncovalent complex is required for processive BRCA1-directed ubiquitination. *Molecular Cell* **21**, 873–880 (2006).
93. Kumar, P. *et al.* Role of a non-canonical surface of Rad6 in ubiquitin conjugating activity. *Nucleic Acids Research* gkv845 (2015). doi:10.1093/nar/gkv845
94. Lange, O. F. *et al.* Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* **320**, 1471–1475 (2008).
95. Ye, Y. & Rape, M. Building ubiquitin chains: E2 enzymes at work. *Nature reviews. Molecular cell biology* **10**, 755–764 (2009).
96. Wenzel, D. M., Stoll, K. E. & Klevit, R. E. E2s: structurally economical and functionally replete. *Biochem. J.* **433**, 31–42 (2011).
97. van Wijk, S. J. L. & Timmers, H. T. M. The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins. *FASEB J.* **24**, 981–993 (2010).
98. Wu, P.-Y. *et al.* A conserved catalytic residue in the ubiquitin-conjugating enzyme family. *EMBO J.* **22**, 5241–5250 (2003).

99. Berndsen, C. E., Wiener, R., Yu, I. W., Ringel, A. E. & Wolberger, C. A conserved asparagine has a structural role in ubiquitin-conjugating enzymes. *Nat. Chem. Biol.* **9**, 154–156 (2013).
100. Cook, B. W. & Shaw, G. S. Architecture of the catalytic HPN motif is conserved in all E2 conjugating enzymes. *Biochem. J.* **445**, 167–174 (2012).
101. Komander, D., Clague, M. J. & Urbé, S. Breaking the chains: structure and function of the deubiquitinases. *Nature reviews. Molecular cell biology* **10**, 550–563 (2009).
102. Komander, D. & Rape, M. The ubiquitin code. *Annu. Rev. Biochem.* **81**, 203–229 (2012).
103. Sowa, M. E., Bennett, E. J., Gygi, S. P. & Harper, J. W. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**, 389–403 (2009).
104. Wolberger, C. Mechanisms for regulating deubiquitinating enzymes. *Protein Sci.* **23**, 344–353 (2014).
105. Wiener, R., Zhang, X., Wang, T. & Wolberger, C. The mechanism of OTUB1-mediated inhibition of ubiquitination. *Nature* **483**, 618–622 (2012).
106. Juang, Y.-C. *et al.* OTUB1 co-opts Lys48-linked ubiquitin recognition to suppress E2 enzyme function. *Molecular Cell* **45**, 384–397 (2012).
107. Wiener, R. *et al.* E2 ubiquitin-conjugating enzymes regulate the deubiquitinating activity of OTUB1. *Nat. Struct. Mol. Biol.* **20**, 1033–1039 (2013).
108. Spratt, D. E. & Shaw, G. S. Association of the disordered C-terminus of CDC34 with a catalytically bound ubiquitin. *J. Mol. Biol.* **407**, 425–438 (2011).
109. Schumacher, F.-R., Wilson, G. & Day, C. L. The N-terminal extension of UBE2E ubiquitin-conjugating enzymes limits chain assembly. *J. Mol. Biol.* **425**, 4099–4111 (2013).
110. Banka, P. A., Behera, A. P., Sarkar, S. & Datta, A. B. RING E3-Catalyzed E2 Self-Ubiquitination Attenuates the Activity of Ube2E Ubiquitin-Conjugating Enzymes. *J. Mol. Biol.* **427**, 2290–2304 (2015).
111. Budhidarmo, R., Nakatani, Y. & Day, C. L. RINGs hold the key to ubiquitin transfer. *Trends in biochemical sciences* **37**, 58–65 (2012).
112. Friedman, P. N., Chen, X., Bargonetti, J. & Prives, C. The p53 protein is an unusually shaped tetramer that binds directly to DNA. *Proceedings of the National Academy of Sciences* **90**, 3319–3323 (1993).
113. Rojas-Fernandez, A. *et al.* SUMO chain-induced dimerization activates RNF4. *Molecular Cell* **53**, 880–892 (2014).
114. Hodson, C., Purkiss, A., Miles, J. A. & Walden, H. Structure of the human FANCL RING-Ube2T complex reveals determinants of cognate E3-E2 selection. *Structure* **22**, 337–344 (2014).
115. David, Y. *et al.* E3 ligases determine ubiquitination site and conjugate type by enforcing specificity on E2 enzymes. *J. Biol. Chem.* **286**, 44104–44115 (2011).
116. Chen, B. *et al.* The activity of a human endoplasmic reticulum-associated degradation E3, gp78, requires its Cue domain, RING finger, and an E2-binding site. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 341–346 (2006).
117. Hibbert, R. G., Huang, A., Boelens, R. & Sixma, T. K. E3 ligase Rad18 promotes monoubiquitination rather than ubiquitin chain formation by E2 enzyme Rad6. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5590–5595 (2011).
118. Christensen, D. E., Brzovic, P. S. & Klevit, R. E. E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages. *Nat. Struct. Mol. Biol.* **14**, 941–948 (2007).
119. Kar, G., Keskin, O., Nussinov, R. & Gursoy, A. Human proteome-scale structural modeling of E2-E3 interactions exploiting interface motifs. *J. Proteome Res.* **11**, 1196–1207 (2012).
120. van Wijk, S. J. L. *et al.* A comprehensive framework of E2-RING E3 interactions of the human ubiquitin-proteasome system. *Mol. Syst. Biol.* **5**, 295 (2009).
121. Markson, G. *et al.* Analysis of the human E2 ubiquitin conjugating enzyme protein interaction network. *Genome Res.* **19**, 1905–1911 (2009).
122. Napolitano, L. M., Jaffray, E. G., Hay, R. T. & Meroni, G. Functional interactions between ubiquitin E2 enzymes and TRIM proteins. *Biochem. J.* **434**, 309–319 (2011).
123. Saha, A., Lewis, S., Kleiger, G., Kuhlman, B. & Deshaies, R. J. Essential role for ubiquitin-ubiquitin-conjugating enzyme interaction in ubiquitin discharge from Cdc34 to substrate. *Molecular Cell* **42**, 75–83 (2011).
124. Campbell, S. J. *et al.* Molecular insights into the function of RING finger (RNF)-containing proteins hRNF8 and hRNF168 in Ubc13/Mms2-dependent ubiquitylation. *J. Biol. Chem.* **287**, 23900–23910 (2012).

125. Ozkan, E., Yu, H. & Deisenhofer, J. Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18890–18895 (2005).
126. Dominguez, C. *et al.* Structural model of the UbcH5B/CNOT4 complex revealed by combining NMR, mutagenesis, and docking approaches. *Structure* **12**, 633–644 (2004).
127. Benirschke, R. C. *et al.* Molecular basis for the association of human E4B U box ubiquitin ligase with E2-conjugating enzymes UbcH5c and Ubc4. *Structure* **18**, 955–965 (2010).
128. Wickliffe, K. E., Lorenz, S., Wemmer, D. E., Kuriyan, J. & Rape, M. The mechanism of linkage-specific ubiquitin chain elongation by a single-subunit E2. *Cell* **144**, 769–781 (2011).
129. Pruneda, J. N. *et al.* Structure of an E3:E2~Ub complex reveals an allosteric mechanism shared among RING/U-box ligases. *Molecular Cell* **47**, 933–942 (2012).
130. Plechanová, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. & Hay, R. T. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115–120 (2012).
131. Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. *Nat. Struct. Mol. Biol.* **19**, 876–883 (2012).
132. Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. Essentiality of a non-RING element in priming donor ubiquitin for catalysis by a monomeric E3. *Nat. Struct. Mol. Biol.* **20**, 982–986 (2013).
133. Vittal, V., Stewart, M. D., Brzovic, P. S. & Klevit, R. E. Regulating the Regulators: Recent Revelations in the Control of E3 Ubiquitin Ligases. *J. Biol. Chem.* **290**, 21244–21251 (2015).
134. DaRosa, P. A. *et al.* Allosteric activation of the RNF146 ubiquitin ligase by a poly(ADP-ribosylation) signal. *Nature* (2014). doi:10.1038/nature13826
135. Scott, D. C. *et al.* Structure of a RING E3 trapped in action reveals ligation mechanism for the ubiquitin-like protein NEDD8. *Cell* **157**, 1671–1684 (2014).
136. McGinty, R. K., Henrici, R. C. & Tan, S. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* **514**, 591–596 (2014).
137. Brown, N. G. *et al.* RING E3 mechanism for ubiquitin ligation to a disordered substrate visualized for human anaphase-promoting complex. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 5272–5279 (2015).
138. Chang, L., Zhang, Z., Yang, J., McLaughlin, S. H. & Barford, D. Atomic structure of the APC/C and its mechanism of protein ubiquitination. *Nature* **522**, 450–454 (2015).
139. Mace, P. D. *et al.* Structures of the cIAP2 RING domain reveal conformational changes associated with ubiquitin-conjugating enzyme (E2) recruitment. *J. Biol. Chem.* **283**, 31633–31640 (2008).
140. Bentley, M. L. *et al.* Recognition of UbcH5c and the nucleosome by the Bmi1/Ring1b ubiquitin ligase complex. *EMBO J.* **30**, 3285–3297 (2011).
141. Zhang, L. *et al.* The IDOL-UBE2D complex mediates sterol-dependent degradation of the LDL receptor. *Genes & development* **25**, 1262–1274 (2011).
142. Taherbhoy, A. M., Huang, O. W. & Cochran, A. G. BMI1-RING1B is an autoinhibited RING E3 ubiquitin ligase. *Nat Commun* **6**, 7621 (2015).
143. Yudina, Z. *et al.* RING Dimerization Links Higher-Order Assembly of TRIM5α to Synthesis of K63-Linked Polyubiquitin. *Cell Rep* (2015). doi:10.1016/j.celrep.2015.06.072
144. Bhaskar, V., Basquin, J. & Conti, E. Architecture of the ubiquitylation module of the yeast Ccr4-Not complex. *Structure* **23**, 921–928 (2015).
145. Dou, H. *et al.* Structural basis for autoinhibition and phosphorylation-dependent activation of c-Cbl. *Nat. Struct. Mol. Biol.* **19**, 184–192 (2012).
146. Hatakeyama, S. TRIM proteins and cancer. *Nature reviews. Cancer* **11**, 792–804 (2011).
147. Kane, R. C. *et al.* Bortezomib for the treatment of mantle cell lymphoma. *Clin. Cancer Res.* **13**, 5291–5294 (2007).
148. Kane, R. C., Bross, P. F., Farrell, A. T. & Pazdur, R. Velcade: U.S. FDA approval for the treatment of multiple myeloma progressing on prior therapy. *Oncologist* **8**, 508–513 (2003).
149. Soucy, T. A. *et al.* An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **458**, 732–736 (2009).
150. Brownell, J. E. *et al.* Substrate-assisted inhibition of ubiquitin-like protein-activating enzymes: the NEDD8 E1 inhibitor MLN4924 forms a NEDD8-AMP mimetic in situ. *Molecular Cell* **37**, 102–111 (2010).

151. Ceccarelli, D. F. *et al.* An allosteric inhibitor of the human Cdc34 ubiquitin-conjugating enzyme. *Cell* **145**, 1075–1087 (2011).
152. Huang, H. *et al.* E2 enzyme inhibition by stabilization of a low-affinity interface with ubiquitin. *Nat. Chem. Biol.* **10**, 156–163 (2014).
153. Short, K. M. & Cox, T. C. Subclassification of the RBCC/TRIM superfamily reveals a novel motif necessary for microtubule binding. *J. Biol. Chem.* **281**, 8970–8980 (2006).
154. Meroni, G. & Diez-Roux, G. TRIM/RBCC, a novel class of ‘single protein RING finger’ E3 ubiquitin ligases. *Bioessays* **27**, 1147–1157 (2005).
155. Sardiello, M., Cairo, S., Fontanella, B., Ballabio, A. & Meroni, G. Genomic analysis of the TRIM family reveals two groups of genes with distinct evolutionary properties. *BMC Evol. Biol.* **8**, 225 (2008).
156. Diaz-Griffero, F. *et al.* A B-box 2 surface patch important for TRIM5 α self-association, capsid binding avidity, and retrovirus restriction. *J. Virol.* **83**, 10737–10751 (2009).
157. Ganser-Pornillos, B. K. *et al.* Hexagonal assembly of a restricting TRIM5 α protein. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 534–539 (2011).
158. Oke, V. & Wahren-Herlenius, M. The immunobiology of Ro52 (TRIM21) in autoimmunity: a critical review. *J. Autoimmun.* **39**, 77–82 (2012).
159. Kong, H. J. *et al.* Cutting edge: autoantigen Ro52 is an interferon inducible E3 ligase that ubiquitinates IRF-8 and enhances cytokine expression in macrophages. *J. Immunol.* **179**, 26–30 (2007).
160. Higgs, R. *et al.* The E3 ubiquitin ligase Ro52 negatively regulates IFN- β production post-pathogen recognition by polyubiquitin-mediated degradation of IRF3. *J. Immunol.* **181**, 1780–1786 (2008).
161. Espinosa, A. *et al.* Loss of the lupus autoantigen Ro52/Trim21 induces tissue inflammation and systemic autoimmunity by dysregulating the IL-23-Th17 pathway. *J. Exp. Med.* **206**, 1661–1671 (2009).
162. Yang, K. *et al.* TRIM21 is essential to sustain IFN regulatory factor 3 activation during antiviral response. *J. Immunol.* **182**, 3782–3792 (2009).
163. Bernatsky, S. *et al.* Cancer risk in systemic lupus: an updated international multi-centre cohort study. *J. Autoimmun.* **42**, 130–135 (2013).
164. Voulgarelis, M. & Skopouli, F. N. Clinical, immunologic, and molecular factors predicting lymphoma development in Sjogren's syndrome patients. *Clin Rev Allergy Immunol* **32**, 265–274 (2007).
165. Sabile, A. *et al.* Regulation of p27 degradation and S-phase progression by Ro52 RING finger protein. *Molecular and cellular biology* **26**, 5994–6004 (2006).
166. Higgs, R. *et al.* The E3 ubiquitin ligase Ro52 negatively regulates IFN- β production post-pathogen recognition by polyubiquitin-mediated degradation of IRF3. *J. Immunol.* **181**, 1780–1786 (2008).
167. Stacey, K. B., Breen, E. & Jefferies, C. A. Tyrosine phosphorylation of the E3 ubiquitin ligase TRIM21 positively regulates interaction with IRF3 and hence TRIM21 activity. *PLoS ONE* **7**, e34041 (2012).
168. Lazzari, E. *et al.* TRIPartite motif 21 (TRIM21) differentially regulates the stability of interferon regulatory factor 5 (IRF5) isoforms. *PLoS ONE* **9**, e103609 (2014).
169. Higgs, R. *et al.* Self protection from anti-viral responses--Ro52 promotes degradation of the transcription factor IRF7 downstream of the viral Toll-Like receptors. *PLoS ONE* **5**, e11776 (2010).
170. Kim, J. Y. & Ozato, K. The sequestosome 1/p62 attenuates cytokine gene expression in activated macrophages by inhibiting IFN regulatory factor 8 and TNF receptor-associated factor 6/NF- κ B activity. *J. Immunol.* **182**, 2131–2140 (2009).
171. Grandinetti, K. B. *et al.* Overexpression of TRIB2 in human lung cancers contributes to tumorigenesis through downregulation of C/EBP α . *Oncogene* **30**, 3328–3335 (2011).
172. Wada, K., Niida, M., Tanaka, M. & Kamitani, T. Ro52-mediated monoubiquitination of IKK {beta} down-regulates NF- κ B signalling. *J. Biochem.* **146**, 821–832 (2009).
173. Niida, M., Tanaka, M. & Kamitani, T. Downregulation of active IKK beta by Ro52-mediated autophagy. *Mol. Immunol.* **47**, 2378–2387 (2010).
174. Zhang, Z. *et al.* The E3 ubiquitin ligase TRIM21 negatively regulates the innate immune response to intracellular double-stranded DNA. *Nat. Immunol.* **14**, 172–178 (2013).

175. Reddy, B. A. *et al.* Nucleotide biosynthetic enzyme GMP synthase is a TRIM21-controlled relay of p53 stabilization. *Molecular Cell* **53**, 458–470 (2014).
176. Plafker, S. M., Plafker, K. S., Weissman, A. M. & Macara, I. G. Ubiquitin charging of human class III ubiquitin-conjugating enzymes triggers their nuclear import. *The Journal of cell biology* **167**, 649–659 (2004).
177. Napolitano, L. M. & Meroni, G. TRIM family: Pleiotropy and diversification through homomultimer and heteromultimer formation. *IUBMB Life* **64**, 64–71 (2012).
178. Yamauchi, K., Wada, K., Tanji, K., Tanaka, M. & Kamitani, T. Ubiquitination of E3 ubiquitin ligase TRIM5 alpha and its potential role. *The FEBS journal* **275**, 1540–1555 (2008).
179. McEwan, W. A. *et al.* Intracellular antibody-bound pathogens stimulate immune signaling via the Fc receptor TRIM21. *Nat. Immunol.* **14**, 327–336 (2013).
180. Tatham, M. H., Plechanovová, A., Jaffray, E. G., Salmen, H. & Hay, R. T. Ube2W conjugates ubiquitin to α -amino groups of protein N-termini. *Biochem. J.* **453**, 137–145 (2013).
181. Fletcher, A. J., Mallery, D. L., Watkinson, R. E., Dickson, C. F. & James, L. C. Sequential ubiquitination and deubiquitination enzymes synchronize the dual sensor and effector functions of TRIM21. *Proc. Natl. Acad. Sci. U.S.A.* 201507534 (2015). doi:10.1073/pnas.1507534112
182. Mallery, D. L. *et al.* Antibodies mediate intracellular immunity through tripartite motif-containing 21 (TRIM21). *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19985–19990 (2010).
183. Sanchez, J. G. *et al.* The tripartite motif coiled-coil is an elongated antiparallel hairpin dimer. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2494–2499 (2014).
184. Streich, F. C., Ronchi, V. P., Connick, J. P. & Haas, A. L. Tripartite motif ligases catalyze polyubiquitin chain formation through a cooperative allosteric mechanism. *J. Biol. Chem.* **288**, 8209–8221 (2013).
185. Li, Y. *et al.* Structural insights into the TRIM family of ubiquitin E3 ligases. *Cell Res.* **24**, 762–765 (2014).
186. Goldstone, D. C. *et al.* Structural studies of postentry restriction factors reveal antiparallel dimers that enable avid binding to the HIV-1 capsid lattice. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9609–9614 (2014).
187. Weinert, C., Morger, D., Djekic, A., Grütter, M. G. & Mittl, P. R. E. Crystal structure of TRIM20 C-terminal coiled-coil/B30.2 fragment: implications for the recognition of higher order oligomers. *Sci Rep* **5**, 10819 (2015).
188. Sastri, J. *et al.* Restriction of HIV-1 by rhesus TRIM5 α is governed by alpha helices in the Linker2 region. *J. Virol.* **88**, 8911–8923 (2014).
189. Li, X., Yeung, D. F., Fiegen, A. M. & Sodroski, J. Determinants of the higher order association of the restriction factor TRIM5alpha and other tripartite motif (TRIM) proteins. *J. Biol. Chem.* **286**, 27959–27970 (2011).
190. Du, H. *et al.* The MID1 E3 ligase catalyzes the polyubiquitination of Alpha4 (α 4), a regulatory subunit of protein phosphatase 2A (PP2A): novel insights into MID1-mediated regulation of PP2A. *J. Biol. Chem.* **288**, 21341–21350 (2013).

Papers

The articles associated with this thesis have been removed for copyright reasons. For more details about these see:

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-122468>