Mechanism of regulation of the p300/CBP acetyltransferase

Manuela Delvecchio

▶ **To cite this version:**

Manuela Delvecchio. Mechanism of regulation of the p300/CBP acetyltransferase. Molecular biology. Université de Grenoble, 2011. English. NNT : 2011GRENV039 . tel-01059819

HAL Id: tel-01059819
https://theses.hal.science/tel-01059819

Submitted on 2 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE GRENOBLE

THESIS

To obtain the title of

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Specialty: **Structural Biology and Nanobiology**

Arrêté ministériel : 7 août 2006

Defended in public by

Manuela DELVECCHIO

Thesis directed by **Daniel PANNE**

prepared at the **European Molecular Biology Laboratory (EMBL)**

within the **École Doctorale Chimie et Sciences du Vivant**

Mechanism of regulation of the p300/CBP acetyltransferase

Defended in public

the **26th September 2011**,

in front of the jury :

Dr. Saadi KHOCHBIN	**Chair**
Dr. Rein AASLAND	**Reviewer**
Dr. Steve GAMBLIN	**Reviewer**
Dr. Christoph MÜLLER	**Reviewer**
Dr. Carlo PETOSA	**Examiner**

European Molecular Biology Laboratory
Grenoble Outstation

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biologie structurale et nanobiologie**

Arrêté ministériel : 7 août 2006

Présentée par

Manuela DELVECCHIO

Thèse dirigée par **Daniel PANNE**

préparée au sein du **Laboratoire Européen de Biologie Moléculaire (EMBL)**

dans **l'École Doctorale Chimie et Sciences du Vivant**

Mécanisme de régulation de l'acétyltransférase p300/CBP

Thèse soutenue publiquement

le **26 Septembre 2011**,

devant le jury composé de :

Dr. Saadi KHOCHBIN	**Président**
Dr. Rein AASLAND	**Rapporteur**
Dr. Steve GAMBLIN	**Rapporteur**
Dr. Christoph MÜLLER	**Rapporteur**
Dr. Carlo PETOSA	**Examinateur**

ABSTRACT

The p300/CBP acetyltransferase is an important transcriptional co-activator which is involved in regulating a wide range of biological processes, such as DNA transcription, development and innate immunity. To date, the role of p300/CBP in gene regulation has been extensively described but little is known about the mechanisms which regulate its activity. Since misregulation of p300/CBP has been associated to the development of several forms of cancers and neurodegenerative diseases, studies directed to decipher the mechanisms of regulation of p300/CBP are of great importance for the development of new therapies.

The p300/CBP 'core' contains two post-translational modifications (PTMs) recognition motifs, a bromodomain and a PHD domain (the bromo-PHD module, BP), in close proximity to a histone acetyltransferase domain (HAT). Many chromatin modifying enzymes contain recognition modules for PTMs. Frequently particular groupings of such modules are conserved and linked within the same protein or the same multisubunit complex, suggesting that they perform concerted functions. These linked modules may act combinatorially to allow recognition of multiple PTMs or display new functions that are not possessed by the single modules, such as regulatory properties.

Accumulating evidence suggests that acetylation/deacetylation in a conserved autoinhibitory loop of the p300/CBP HAT domain plays an important role in regulation of HAT activity. The close apposition of the BP module and the HAT domain suggests that BP substrate recognition is intrinsically linked to regulation of HAT activity.

During my thesis work, I have investigated the role of BP in HAT regulation. I propose that the BP module is involved in p300/CBP regulation by binding to the HAT domain and stabilizing the autoinhibited conformation of the enzyme. I have also investigated substrate specificity of the BP module towards modified chromatin. I could show that the BP module binds histone modifications including those that are p300/CBP dependent. Altogether, the data suggests that the BP module is involved in regulating p300/CBP HAT activity and in targeting of chromatin.

RÉSUMÉ EN FRANÇAIS

Le p300/CBP acétyltransférase est un co-activateur transcriptionnel très important qui est impliqué dans la régulation d'un grand nombre de processus biologiques, comme la transcription d'ADN, le développement et l'immunité innée. Jusqu'à présent, le rôle de p300/CBP dans la régulation de l'expression des gènes a été largement étudiée, mais les mécanismes qui régulent son activité enzymatique sont encore peu connus. Des études ont montré que le dysfonctionnement de p300/CBP est associé à plusieurs formes de cancer et de maladies neurodégénératives. Dés lors, chaque progrès concernant les mécanismes de régulation de p300/CBP est devenu primordial pour le développement de nouvelles thérapies.

Le 'noyau' de p300/CBP contient deux domaines pour la reconnaissance des modifications post-traductionnelles (MPTs), un bromodomaine et un PHD *finger* (le module BP), adjacent à un domaine HAT (ou domaine histone acétyltransférase). Plusieurs enzymes, modifiant la chromatine, contiennent des domaines de reconnaissance des MPTs. Fréquemment des groupements particuliers de ces domaines sont très conservés et liés, au sein de la même protéine ou du même complexe protéique, suggérant qu'ils réalisent des fonctions coordonnées. Ces domaines adjacents peuvent agir en concertation dans la reconnaissance simultanée de différents MPTs ou peuvent exercer des fonctions différentes de celles qui sont effectuées par ces deux domaines particuliers, tels que les fonctions de régulation enzymatique.

Plusieurs études suggèrent que les cycles acétylation/désacétylation dans la boucle d'auto-inhibition, à l'intérieur du domaine HAT, jouent un rôle important dans la régulation de l'activité enzymatique de p300/CBP. La proximité du module BP et du domaine HAT suggère que la spécificité de liaison, appartenant au module BP, peut être intrinsèquement liée à la régulation de l'activité du domaine HAT.

L'objectif de ma thèse est de déterminer le rôle du module BP dans la régulation de l'activité du domaine HAT. Je propose que le module BP soit impliqué dans la régulation de p300/CBP de deux façons. La première consiste à établir un lien avec le domaine HAT qui stabilise la conformation auto-inhibée de l'enzyme. La deuxième exige que le module BP

joue un rôle dans le choix des substrats de p300/CBP. J'ai été en mesure de montrer que BP peut se lier au domaine HAT et à la chromatine modifiée et qu'il peut reconnaître les modifications effectuées par p300/CBP lui-même. Les données obtenues indiquent que le module BP peut être impliqué dans la régulation de l'activité de p300/CBP et dans son ciblage à la chromatine.

LIST OF ABBREVIATIONS

A	aa	amino acid
	Abs	Absorbance
	AcCOA	Acetyl-coenzyme A
	ADP	Adenosine diphosphate
	AT	Acetyl-Transferase
	ATP	Adenosine-5'-triphosphate
B	bp	base pair
	BP	Bromo-PHD module
	BPTF	Bromodomain and PHD domain Transcription Factor
	BRD	Bromodomain
	Brdt-BD1	First bromodomain of Brdt, a member of the BET sub-family
	BSA	Bovine Serum Albumin
C	CBP	CREB Binding Protein
	CCD	Charge-Coupled Device
	CH	Cysteine-Histidine rich domain
	CHD	Chromo-Helicase/ATPase DNA binding
	ChIP	Chromatin Immunoprecipitation
	COA(SH)	Coenzyme A
	CREB	cAMP Response Element-Binding
	CyP33	Cyclophilin 33
D	DLS	Dynamic Light Scattering
	DNA	Deoxyribonucleic Acid
	DPF3	D4, zinc and double PHD fingers, family 3
	DTNB	Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid)
	DTT	Dithiothreitol
	Da	Dalton

E	EDTA	Ethylenediaminetetraacetic Acid
	EMSA	Electrophoretic Mobility Shift Assay
	ESA1	Essential SAS2-related acetyltransferase
	ESRF	European Synchrotron Radiation Facility
F	FAT	Factor Acetyl-Transferase
	FT	Flow Through
G	Gcn5	General control of amino-acid synthesis 5
	GFP	Green Fluorescent Protein
	GNAT	Gcn5-related N-Acetyl-Transferases
	GST	Glutathione-S-transferase
	GTP	Guanosine-5'-triphosphate
H	HA	Haemagglutinin
	HAT	Histone Acetyl-Transferase
	HDAC	Histone Deacetylase
	Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
	HFD	Histone Fold Domain
	His-tag	Polyhistidine tag
	HMT	Histone Methyl-Transferase
	HRP	Horseradish Peroxidase
I	IBiD	IRF3-Binding Domain
	IFN	Interferon
	IgG	Immunoglobulin G
	Ino80	DNA helicase (= YGL150C)
	IPTG	Isopropyl-β-D-thio-galactoside
	IRF	Interferon Regulatory Factor
	ISWI	Imitation SWI of *Drosophila melanogaster*
	ITC	Isothermal Titration Calorimetry
J, K	KAP1	KRAB Associated Protein 1
	KAT	Lysine (K)-Acetyl-Transferase
	KDM	Lysine (K)-Demethylase
	KIX	CREB binding domain

	KMT	Lysine (K)-Methyl-Transferase
L	LB	Luria Broth
	LPS	Lipopolysaccharide
M	MALLS	Multi-Angle Laser Light Scattering
	MBT	Malignant Brain Tumour
	MCS	Multiple Cloning Site
	MLL1	Mixed Lineage Leukaemia 1
	MOF	Male absent on the first
	MS	Mass Spectrometry
	MyoD	Myoblast Determination protein 1
	MYST	MOZ, Ybf/Sas3, Sas2, Tip60
N	NAD	Nicotinamide Adenine Dinucleotide
	NCS	Non-Crystallographic Simmetry
	NDR	Nucleosome Depleted Region
	NMC	NUT Midline Carcinoma
	NMR	Nuclear Magnetic Resonance
	NUT	Nuclear protein in Testis
O	OD	Optical Density
	o/n	Overnight
P	P21	protein 21 or cyclin dependent kinase inhibitor 1
	P53	protein 53
	P300	protein of 300 kDa or E1a binding protein
	PAMP	Pathogen Associated Molecular Pattern
	PCAF	p300/CBP Associated Factor
	PDB	Protein Data Bank
	PEG	Polyethylene glycol
	PHD	Plant Homeodomain
	PIC	Pre-Initiation Complex
	PRR	Pattern Recognition Receptor
	PTM	Post Translational Modification

Q, R	RING	Really Interesting New Gene
	RNA	Ribonucleic Acid
	Rpm	Revolutions per minute
	RRM	RNA Recognition Motif
	RT	Room Temperature
	Rtt109	Regulation of Ty1 transposition protein 109
	RSC	Remodels the Structure of Chromatin
S	S200	Superdex 200
	SANT	SWI3-ADA2-N-CoR-TFIIIB
	SAXS	Small Angle X-ray Scattering
	SDS	Sodium Dodecyl Sulphate
	SeMet	Seleno Methionine
	SH2	Src Homology 2
	Sir2	Silent information regulator-2
	SIRT2	Human Sirtuin 2
	SWI/SNF	SWItch/Sucrose Non Fermentable
T	TAF1	TBP Associated Factor 1
	TCEP	Tris(2-carboxyethyl)phosphine
	TEV	Tobacco Etch Virus
	TF	Transcription Factor
	Tris	Tris(hydroxymethyl)aminomethane
	TSS	Transcriptional Start Site
	TATA box	Sequence of 7 bases (TATAAAA) found at the promoter region
U, V, W	WD40	β-transducin repeat
	UV	Ultraviolet
X, Y, Z	ZYP 5052	Auto-inducing medium

TABLE OF CONTENTS

1. INTRODUCTION

RÉSUMÉ EN FRANÇAIS

p300/CBP est un co-activateur transcriptionnel et une enzyme de modification post-traductionnelle des histones appartenant à la classe des acétyltransférases. La modification post-traductionnelle des histones est l'une des stratégies adoptées par les acteurs de la transcription pour surmonter la barrière imposée par la chromatine qui compacte le génome en limitant l'accès aux régions codantes. L'organisation du génome, la structure de la chromatine et les stratégies utilisées par la machinerie de transcription pour accéder à des régions codantes du génome sont décrites dans le chapitre 1.2. Le chapitre 1.3 couvre les différents types de modifications post-traductionnelles présentes sur les queues N-terminales des histones ainsi que les domaines qui reconnaissent ces modifications.

L'acétylation des lysines est l'une des modifications post-traductionnelles les plus caractérisées. Un des mécanismes de contrôle de l'expression des gènes mis en œuvre dans plusieurs processus cellulaires est de maintenir un équilibre entre les niveaux d'acétylation et de désacétylation des histones et des facteurs transcriptionnels. En général, cette balance est maintenue grâce à une équilibration de l'activité des acétyltransférases et des désacétylases qui peuvent communiquer entre elles et mutuellement affecter leurs activités. Un exemple d'intercommunication est celle qui se produit entre une désacétylase appelée SIRT2 et l'acétyltransférase p300/CBP. Acétyltransférases et désacétylases sont décrites dans le chapitre 1.4.

On connaît beaucoup de choses sur le rôle joué par l'acétyltransférase p300/CBP dans la régulation des gènes, mais il n'y a que peu d'études sur les mécanismes qui régulent son activité. Le chapitre 1.5 contient les informations connues à ce jour sur les fonctions, la biochimie et la structure de p300/CBP.

Le chapitre 1.6 décrit l'approche utilisée pour prouver mon hypothèse sur le mécanisme de régulation de p300/CBP. Je propose l'idée qu'un module situé à l'intérieur de la région catalytique de p300/CBP peut contribuer de manière significative à la régulation enzymatique de cette dernière. Ce module est constitué de deux domaines souvent associés à la reconnaissance des modifications post-traductionnelles: un bromodomaine et un domaine doigt de zinc spécifique, le PHD *finger*.

1. INTRODUCTION

1.1. IMPORTANCE OF THE STUDY AND THESIS OVERVIEW

Regulation of transcription is the central mechanism by which a cell responds to environmental stimuli. Events which alter this highly regulated process, such as genetic alterations in genes encoding regulatory proteins, can result in severe disorders such as developmental disorders or cancer [1, 2].

The p300/CBP acetyltransferase is a transcriptional co-activator which is thought to be involved in regulation of most eukaryotic genes. It functions as an integrator of a wide variety of signaling pathways and it is utilized by several DNA-binding proteins to facilitate transcriptional activation. P300/CBP plays pivotal roles in regulating many biological processes such as DNA transcription, DNA repair, development, apoptosis and innate immune response. Therefore, it is not surprising that misregulation of p300/CBP has been associated with the development of several forms of cancer and neurodegenerative diseases [2-5].

However, very little is known about the mechanisms that regulate p300/CBP activity [6, 7]. Studies performed in this direction are of great importance, not only to understand the molecular mechanisms that govern the appearance of severe disorders, but also to yield important knowledge about new therapies and strategies to prevent effectively cancer and other diseases.

The goal of my thesis work was to biochemically and structurally characterize the catalytic 'core' of p300/CBP in order to shed light on the mechanisms of protein regulation. Accumulating evidence suggest that cycles of acetylation and deacetylation events regulate p300/CBP activity; in a similar way to phosphorylation and dephosphorylation events in regulating kinases [6, 8]. My hypothesis was that effector modules within the p300/CBP catalytic 'core' are involved in activity regulation. These effector modules comprise a bromodomain (acetyl-lysine recognition module) [9] and a PHD finger (putative methyl-lysine recognition module) [10].

As aforementioned, p300/CBP is a transcriptional co-activator and a histone modifying enzyme which belongs to the family of acetyltransferases. Post-translational

modification and remodeling of histones are some strategies utilized by the transcriptional apparatus to overcome the physical barrier imposed by the nucleosomal structure and to have access to the coding regions of the genome. In part 1.2, I will describe the chromosome structure and the ways by which the transcriptional apparatus overcomes this physical barrier.

Post-translational modifications (PTMs) will be described in part 1.3, as well as the effector modules which recognize them. PTMs can target histone and non-histone proteins, which are both crucial in cellular mechanisms. The presence of modules which recognize PTMs is a common feature in signaling proteins, which often use PTMs as docking site for protein targeting and/or activity.

Amongst these modifications, lysine acetylation is one of the best characterized. Histone acetylation is one of the first chromatin modifications found to be associated with active chromatin [11]. The current model is that histone acetyltransferases (HATs) and deacetylases (HDACs) dynamically regulate the level of chromatin acetylation. For instance, in the case of the p300/CBP acetyltransferase, the deacetylases SIRT2 is thought to counterbalance the activity of p300/CBP. In part 1.4, I will describe HATs and HDACs.

The p300/CBP acetyltransferase is central to transcriptional regulation. However, its mechanism of action and mode of regulation are poorly characterized. In part 1.5, I will describe what is structurally and functionally known about p300/CBP.

Finally, in part 1.6, I will describe the approach followed to demonstrate my

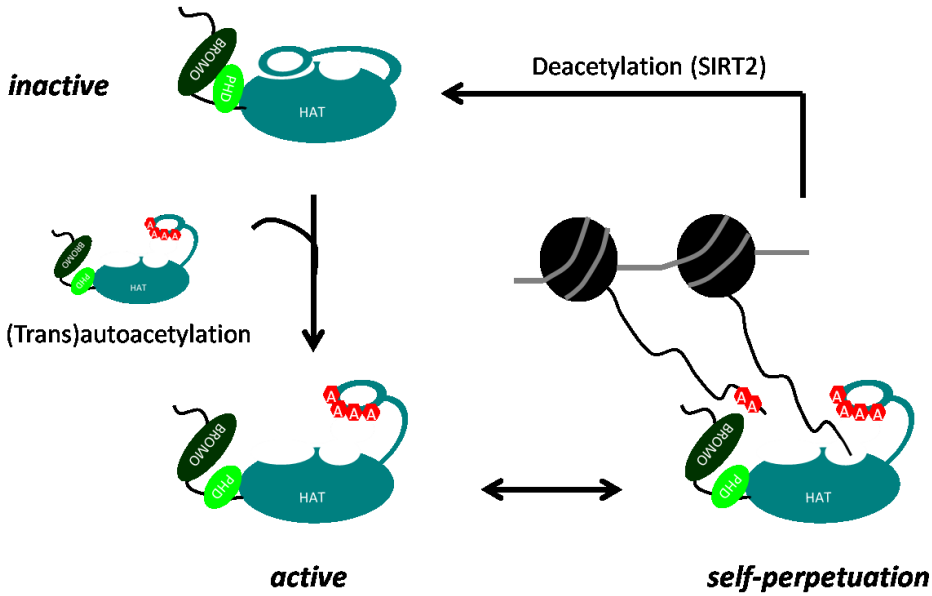


Figure 1 **Proposed model for p300/CBP 'core' regulation.** Cycles of acetylation/deacetylation events are important for p300/CBP activity, as well as bromo- and PHD- domains. In the figure are shown the bromodomain (bottle green), the PHD domain (apple green) and the HAT domain (navy green). Acetylated lysines are indicated by red hexagons. Nucleosomes: black circles; histone tail: black line; DNA: gray line.

working hypothesis. My hypothesis is that the effector modules (bromodomain and PHD finger) are not only involved in substrate binding but also in the regulation of p300/CBP **(Figure 1)**.

1.2. CHROMATIN STRUCTURE AND REMODELING

1.2.1. Gene expression and chromatin

Cells need to respond rapidly to stimuli from the external environment. This often results in rapid regulation of gene transcription upon signaling cascade events, which leads to activation of inducible genes and production of new proteins [1, 12]. Inducible genes are highly regulated genes which must be quickly and specifically activated in response to external stimuli and rapidly inactivated once the stimuli are removed. Well-known examples of inducible genes in higher eukaryotes are the ones implicated in the production of the human interferon α and β (Type I family of interferons, IFN α/β) [13, 14].

Following an infection, expression of type I interferons is induced after recognition of pathogen-associated molecular patterns (PAMPs) by various host pathogen recognition receptors (PRRs). The final step of the signalling cascade is usually the recruitment of the p300/CBP acetyltransferase to the interferon α/β enhanceosome and the activation of gene transcription [14].

In general, gene activation is induced by sequential recruitment of large multi-subunit co-activator complexes which bind to transcriptional activators [12]. This recruitment can be hindered by the compaction of the eukaryotic genome in the nucleus in form of chromatin [15]. Chromatin can represent a physical barrier for the access to DNA sequences and coding regions. Transcriptional activators overcome the problem by recruiting remodeling complexes and enzymes, such as the p300/CBP acetyltransferase, which can alter the chromatin environment and allow access to the DNA [16]. The presence of these chromatin remodelling factors bound to activators facilitates the recruitment of the components of the basal transcriptional machinery to the target genes and thus the rapid activation of transcription [12, 16].

1.2.2. Chromatin structure

The eukaryotic genome is composed of long filaments of DNA. It was calculated that a human genome, fully extended, would measure about a meter [1]. These DNA filaments need to be compacted in a complex structure called chromatin to be able to fit into the nucleus, which measures roughly 10 μm in diameter [1].

Chromatin is an organized combination of DNA and proteins and allows multiple levels of DNA compaction within the cell. The purpose of chromatin organization is not only to package the DNA into a smaller volume to fit in the cell but also to strengthen the DNA to allow mitosis and meiosis. It serves to control vital processes like DNA replication, transcription and repair [17, 18].

There are three basic levels of chromatin compaction **(Figure 2)**: DNA wrapping around nucleosomes to form a beads on a string structure (primary structure of chromatin), nucleosomes interacting with each other to form a 30 nm condensed chromatin fiber (secondary structure of chromatin), chromatin fibers compacted in coils and DNA loops attached to nuclear scaffolds to form chromosomes (the higher order structures) [19].

The precise structure of the 30 nm fiber is currently under investigation, as there are different studies which support alternate architectures [20-24]. There are a number of reviews which discuss the controversy about the 30 nm fiber [25, 26]. The lack of definitive information on chromatin secondary structure complicates the understanding of higher order levels of chromatin organization. Nevertheless, there are proposed models describing possible intermediate levels of chromatin compaction up to the final structure of compacted chromatin (the mitotic chromosome) [27-29] **(Figure 2)**.

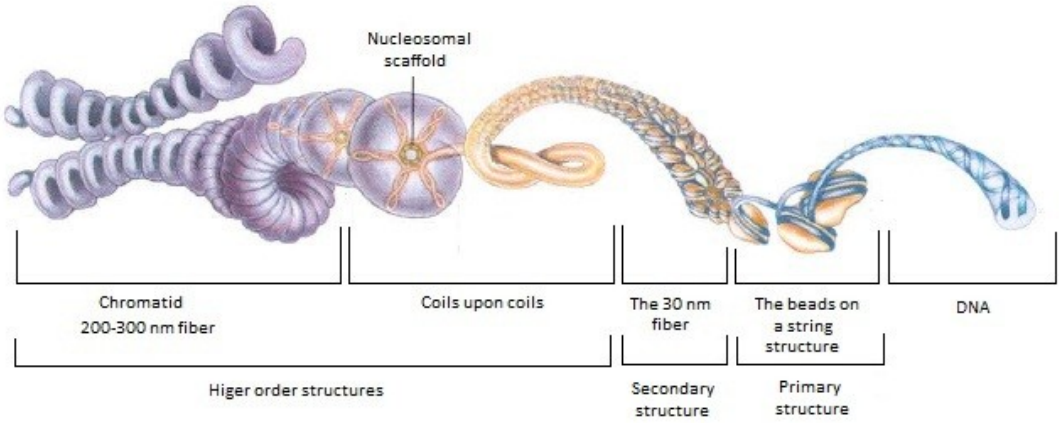


Figure 2 **DNA compaction into chromatin.** DNA compaction in eukaryotic chromosomes involves several levels of organization and the formation of coils upon coils. However, the organization of higher order structures (above 30 nm fibers) in the cell is likely to vary between cells, within region of chromosomes or with time and to be less uniform than depicted here. Figure adapted from [1].

1.2.2.1. The Histones

The DNA in the chromatin is tightly associated with proteins called histones. The histone proteins package and fold the DNA into structural units called nucleosomes. In addition to the compaction of DNA, histones interact with each other and with other distinct chromosomal proteins. These interactions are common in all eukaryotes and contribute to the high degree of conservation in histone sequence through evolution [30].

In eukaryotes there are five major classes of histones, differing in molecular weight and amino acid composition. The histones H2A (14 kDa), H2B (14 kDa) H3 (15 kDa) and H4 (11 kDa) are called core histones, whereas the histone H1 (21 kDa) is called linker histone. Moreover, the aminoacid sequences of histones H3 and H4 are nearly identical in all eukaryotes, whereas histones H2A, H2B and H1 show less sequence similarity across species [31].

Overall, the core histones are small, basic proteins containing relatively large amounts of lysine (Lys) and arginine (Arg) residues. All four core histones contain a globular domain and a charged amino (N)-terminal tail [31]. The globular domain comprises the histone fold domain (composed of a symmetric duplication of helix-strand-helix motifs) **(Figure 3)** [32], some accessory helices and other less structured regions. Histone-histone and histone-DNA interactions occur through the globular domain. The N-terminal tail contains the bulk of the lysine and arginine residues and is the site for many post-translational modifications (PTMs) [31]. The amino acid sequence of the charged tails and the histone fold domains are highly conserved.

The linker histone H1 represents another family of histones and does not have the same structure as the core histones [33]. Linker histones are highly basic, being particularly rich in lysines, and slightly larger than the core histones. They have a central globular domain of about 80 amino acids **(Figure 3)** and highly charged tails at both the amino (N)-

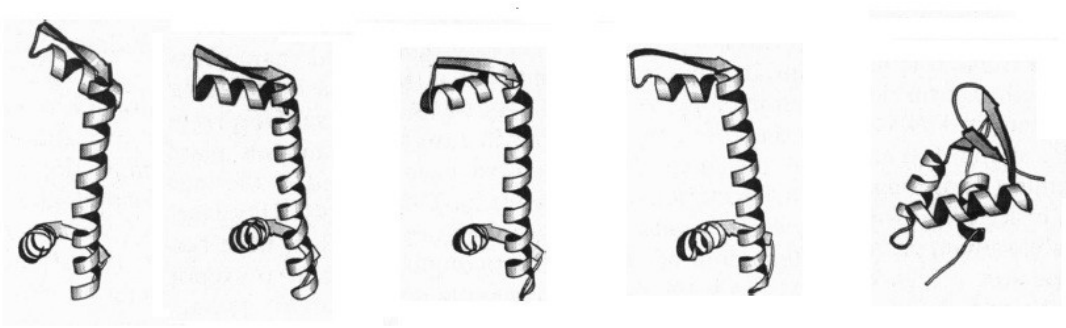


Figure 3 **The histone fold domain.** From left to right, the structures of the histone fold domain of H2A, H2B, H4 and H3. For comparison, the globular core structure of histone H1 is shown on the right [32].

and carboxyl (C)-termini. These N- and C- terminal tails diverge in size and sequence among different H1 variants. Linker histones are not found within the nucleosome core particle. Instead, H1 histones exhibit high affinity for DNA only after it is incorporated into the nucleosome structure. In particular, they bind at the entry and exit sites of the DNA into the nucleosome core particle. This stabilizes the interaction between the DNA and the core histones and forms the so-called chromatosome (nucleosomal structure associated with H1). Stabilization of the nucleosomal structure is important for the folding of nucleosomal arrays in the chromatin fiber [30, 33].

1.2.2.2. The nucleosome

The fundamental repeating unit of chromatin is the nucleosome. The nucleosome core particle consists of about 147 base pairs (bp) of DNA wrapped in a 1.67 left-handed helical turn around a histone octamer. The histone octamer is composed of two copies of each of the core histones (an H3/H4 tetramer and two H2A/H2B dimers). Core particles are interconnected by stretches of linker DNA which can be up to 80 bp long [15, 33]. Non-condensed nucleosomes without the linker histone resemble "beads on a string of DNA" under an electron microscope **(Figure 4)** [34].

The first crystal structure of the nucleosome was solved by the Richmond group in 1997 **(Figure 4)** [35]. Overall, the core particle has a disc-like shape, of 11 nm in diameter and 5.6 nm in width. The structure has a pseudodyad axis of symmetry which passes

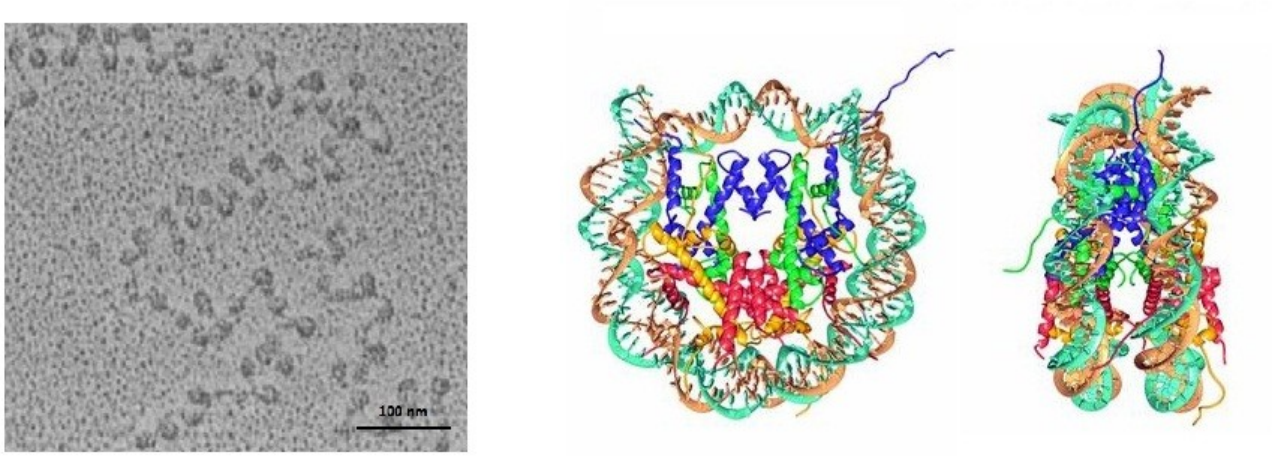


Figure 4 **"Beads on a string" and nucleosome structures.** On the left, non-condensed nucleosomes resemble "beads on a string of DNA" under an electron microscope [34]. On the right, the nucleosome structure as described by Luger *et al.* (1997).

through the centre of the nucleosomal DNA. A tetramer of H3/H4 histones binds to the central 120 bp of the DNA. The two H2A/H2B dimers associate with either side of the H3/H4 tetramer and form a sandwich, extending protein-DNA interaction to the full 147 bp.

Contacts between the histone octamer and DNA are direct or water-mediated. A number of basic and hydroxyl groups as well as side chains amides form salt bridges and hydrogen bonds with the phosphate backbone of the DNA [36]. The distribution and strength of these histone-DNA binding sites makes nucleosomes electrostatically stable but distorts the DNA within the nucleosome core. The overall twist of nucleosomal DNA is only 10.2 bp per turn (varying from a value of 9.4 to 10.9 bp) whereas the one of free B-form DNA in solution is 10.5 bp per turn [37].

The N-terminal tails of the core histones do not directly contribute to the nucleosome core structure. However, the tails are highly positively charged and can contact the phosphate backbone of the DNA or some acidic patches present in other core histones. Work by Dorigo *et al.* showed that higher order folding of the chromatin fiber can be mediated by the histone tails [38].

1.2.3. Chromatin dynamics

Apart from compacting DNA, the chromatin fiber and the nucleosomes also have the important function of regulating the DNA accessibility. Therefore, they exercise direct control over important processes such as DNA transcription, replication and repair.

The need of the cell to respond quickly to external stimuli already suggests that chromatin organisation must be dynamic [39]. There are four principal mechanisms that contribute to chromatin dynamics: control of nucleosome positioning, ATP-dependent chromatin remodeling, replacement of histones with their variants and post-translational modification of histones.

Nucleosome positioning is defined as the probability that a nucleosomes starts at a given base pair within the genome [40]. In general, nucleosomes are more populated in coding regions as compared with intergenic or non-coding regions [41]. Two general types of nucleosome positioning patterns can be present at the promoter region and be classified as open promoters or closed (or occupied) promoters. On one hand, open promoters are generally TATA-less and have a large nucleosome depleted region (NDR, ~150 bp) upstream of the transcriptional start site (TSS) [41]. Such promoters are generally associated with constitutive genes that do not require fine regulation. On the other hand, occupied

promoters do not have an NDR region and contain the TATA box (~25-125 bp from the TSS) which is usually buried inside the nucleosome [41]. These promoters are frequently associated with genes that require tight regulation. In particular, in these promoters the TSS and most of the transcription factors (TFs) binding sites are covered by the nucleosome core. There is only one site (on the linker DNA or the core DNA close to the nucleosome entry) accessible for the so-called "pioneer" transcription factors. The other TF binding sites require chromatin remodelers and histone modifications to become accessible [42].

Nucleosome positioning is also dictated by the length of the linker DNA sequence, which varies strongly between organisms or different tissues and also depends on the transcriptional activity of the genome [41]. In general, long-linker chromatin is less transcriptionally active as compared with short-linker chromatin. Distribution of short linker-chromatin and long-linker chromatin in the nucleus can vary due to the action of chromatin remodelers which are able to slide nucleosomes [43]. For example, Yamada *et al.* showed recently that the remodeling factor ISW1a can set the spacing between two adjacent nucleosomes by acting as a "protein ruler" [44].

ISWIa belongs to a family of ATP-dependent nucleosome remodeling complexes called ISWI. ATP-dependent nucleosome remodeling complexes contain ATPases belonging to the SNF2 subfamily of the nucleic acid stimulated DEAD/H ATPases [43]. They can be subdivided into several families depending on the distinct ATPase domain features. Among these families, the best-known are the SWI2, CHD, Ino80 and the already mentioned ISWI.

In general, ATP-dependent nucleosome remodeling complexes are able to bind DNA, hydrolyze ATP and use the energy to disrupt the nucleosomal structure. This allow a particular region of DNA to become accessible for the binding of specific regulatory factors [45]. There are three different methods that alter DNA accessibility. Remodeling complexes may act by altering the association of the histone molecules within the octamer destabilizing the nucleosome structure (nucleosome remodeling). Secondly, they may move nucleosomes along the DNA to expose a particular DNA sequence (nucleosome sliding). Finally, they may evict a nucleosome from DNA (nucleosome displacement) [46].

An integral part of chromatin dynamics is also the modulation of nucleosomes. There are two ways in which the nucleosome (and therefore the chromatin) structure can be adapted to perform specialized functions: by histone modification and by replacement of histones with their variants. Such changes in nucleosome structure lead to alteration of nucleosome-nucleosome or DNA-histone interactions and this can have important consequences for processes such as DNA transcription and repair.

As aforementioned, there are four conventional core histones: H2A, H2B, H3 and H4. They represent the majority of the histones in the cell (60-90%) and they all have variant subspecies. The histone linker H1 has also several isoforms (eight in total) [47]. Transcriptional activation is only one of the effects caused by the replacement of canonical core histones with their variants. **Figure 5** shows some of the core variants and their proposed function.

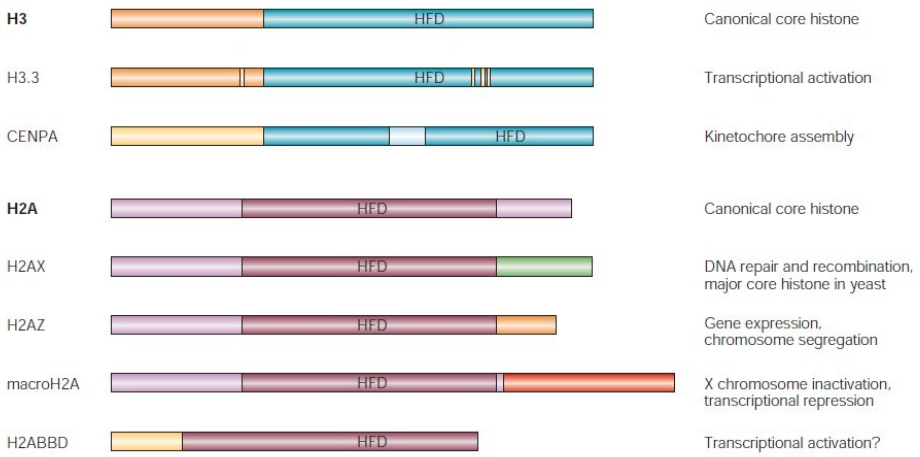


Figure 5 **Histone variants.** Shown here are the core histones H3 and H2A and few of the known variants. HFD denotes the histone fold domain, the structural domain common to all core histones [32]. In histone H3.3 the residues that differ from the major histone H3 are depicted in yellow. The proposed functions of the variants are listed on the right [47].

1.3. HISTONE POST-TRANSLATIONAL MODIFICATIONS

Post-translational modification of histones is one of the major contributions to chromatin plasticity. Covalent modification of the nucleosome can regulate gene expression directly by disrupting chromatin contacts or indirectly by recruiting non-histone proteins to chromatin [48].

Several families of enzymes that modify histones have been identified and most of them are conserved [48]. They mostly target the N-terminal tails of histone proteins, which protrude out of the nucleosome core and are accessible for binding. Accessibility to the histones globular regions is restricted because of structural constraints imposed by the histone fold domain. As a consequence, histone PTMs do not usually affect integrity of the nucleosome structure, as the nucleosome is stabilized by the globular regions of the core histones. Rather, PTMs can have an impact on higher order organization of chromatin, since

they are mostly on N-terminal tails. Few PTMs can be found in the globular domain of core histones, such as the acetylation of lysine 56 on histone H3 by the acetyltransferase p300/CBP [49]. This modification promotes chromatin disassembly during transcriptional activation [49, 50].

There are at least eight distinct types of histone modifications **(Figure 6)** [48]. The best-characterized ones are acetylation, methylation and phosphorylation. Of all the enzymes that modify histones, the methyltransferases (methylation) and the kinases (phosphorylation) are the most specific. In general, acetyltransferases (acetylation) modify more than one lysine residue and tend to have a broader specificity [48].

More than 60 residues on histones can be post-translationally modified. It is difficult to determine the number and the distribution of these PTMs because they are changing rapidly during the cellular life-time and because they depend on the signaling status within the cell. Further complexity is obtained by the fact that the same residue can be targeted by different modifications. For example, lysines can either be acetylated, methylated, ubiquitinated or sumoylated. Methylated lysines can carry one, two or three methyl marks [48]. Moreover, most modifications are dynamic and are rapidly changing due to the presence of modifying enzymes. For example, histone acetyltransferases (HATs or KATs) are responsible for the positioning of acetyl-marks on lysines whereas deacetylases (HDACs) carry out their removal [51]. The same is true for methyltransferases (HMTs or KMTs) and demethylases (KDMs) in the turnover of methylation-marks [52]. For long time, the only example of non-reversible modification was thought to be the methylation of arginine residues. The process only known to remove methylation-marks from arginines was deimination, which converts methyl-arginine in peptidyl-citrulline [53]. However, recently Chang and colleagues described the first arginine demethylase, showing that all known PTMs are reversible [54].

Chromatin Modifications	Residues modified	Functions Regulated
Acetylation	K-ac	Transcription, Repair, Replication, Condensation
Methylation (lysines)	K-me1, K-me3, Kme3	Transcription, Repair
Methylation (arginines)	R-me1, R-me2a, R-me2s	Transcription
Phosphorylation	S-ph, T-ph	Transcription, Repair, Condensation
Ubiquitylation	K-ub	Transcription, Repair
Sumoylation	K-su	Transcription
ADP ribosylation	E-ar	Transcription
Deimination	R > Cit	Transcription
Proline isomeration	P-cis > P-trans	Transcription

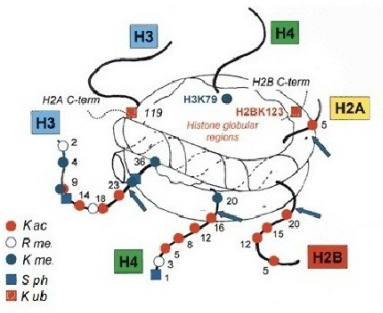


Figure 6 **Different classes of histone modifications.** On the left, an overview of histones modifications. Functions associated with each modification are shown. On the right, a representation of the best studied PTMs. Coloured symbols indicate the sites of PTM and the residue number for each PTM is indicated. Figure adapted from [48].

There are two important mechanisms of action for histone PTMs. The first mechanism is disruption of contacts between nucleosomes or between histones and DNA by acetylation or phosphorylation. The second mechanism is that PTMs serve as signature motifs for recruitment of non-histone proteins into chromatin. Recruitment occurs through chromatin-reader domains that recognize particular PTMs. Frequently, the recruited proteins serve as adaptors which recruit repressor or activator complexes. Similarly important is the presence of histone modifications which prevent docking of non-histone proteins into chromatin [48].

Given the abundance of modifications on histone tails, it is thought that some kind of "combinatorial communication" among PTMs might occur [55]. Firstly, there may be antagonism between distinct types of modification. For example, different modifications on lysine residues are usually mutually exclusive. Secondly, the binding of a protein (or an enzyme) to a specific PTM can be disrupted or strengthened because of adjacent modifications. Thirdly, there could be interplay between modifications which are on different histone tails and/or on different nucleosomes [48, 55].

Each type of PTM can have an important biological function. In the next section, as an example, I will describe the biological significance of lysine acetylation.

1.3.1. The biological function of lysine acetylation

Lysine acetylation is a reversible reaction catalyzed by acetyltransferases and deacetylases. The implication of this PTM in regulation of many different biological processes is reflected by its presence on both histone and non-histone proteins. Its role in regulation can be exerted following two main mechanisms: it can serve as a docking site for non-histone proteins and complexes (which possess specific acetyl-lysine recognition modules, the bromodomains) or it can act by neutralizing the basic charge of lysine residues on histones proteins (which results in alterations of the chromatin structure due to loosening of DNA-histone interactions) [48].

Lysine acetylation can be involved in regulation of gene transcription, DNA repair, DNA replication, chromatin condensation and some very important metabolic processes such as glycolysis and gluconeogenesis [56].

Two very well studied examples indicating distinct biological functions for lysine acetylation are those concerning the acetylations of lysine 16 of histone H4 (H4K16ac) and of lysine 56 of histone H3 (H3K56ac). The H4K16 acetylation has been associated with

transcriptional activation and chromatin decondensation. This modification is catalyzed by several HATs including the p300/CBP acetyltransferase and reversed by a histone deacetylase of the Sir2 family named SIRT2 [57]. The H3K56 acetylation is also catalyzed by p300/CBP and removed by SIRT2. This modification has been linked to gene transcription regulation as well as DNA repair. It is also one of the few examples of PTMs found on the globular domains of core histones [49].

1.3.2. Is there a histone code?

The presence of different patterns of PTMs on histone tails and the possibility to have cross-talk between modifications led to the hypothesis of a "histone code". According to this hypothesis, "multiple histone modifications, acting in a combinatorial or sequential fashion on one or multiple histone tails, specify unique downstream functions" [58]. This would mean that specific combinations of histone tail modifications constitute a code, the histone code, which ultimately dictates biological functions or the state of chromatin. According to this hypothesis the histone code can be "read" by effector proteins that determine specific biological functions. Therefore, the PTMs may be directly predictive of biological outputs.

The presence of cross-talk between modifications and the role of specific modifications in creating binding surfaces for effector molecules which mediate biological outputs were supported by different findings [55]. However, the idea of a code where specific PTMs are predictive of unique downstream functions is controversial. There are several examples showing that the consequences of particular histone modifications are cell context dependent. For example, one of the best studied histone modifications is the trimethylation on lysine 4 of histone H3 (H3K4me3). H3K4me3 is typically associated with active transcription. However, under conditions of DNA damage this modification is used to silence transcription through recruitment of a repressor complex [59]. Such findings show that particular PTMs are not necessarily predictive of biological functions. Hence, it is likely that histone modifications are the result of the current signaling status of the cell, not representing a 'code' per se.

An alternative model to the "histone code" is the "signaling network" hypothesis [60]. According to this hypothesis, histone modifications act as signatures for regulatory mechanisms involving protein-protein interactions, conformational changes and cascades of activity. The direct comparison is with signaling transduction cascades, where

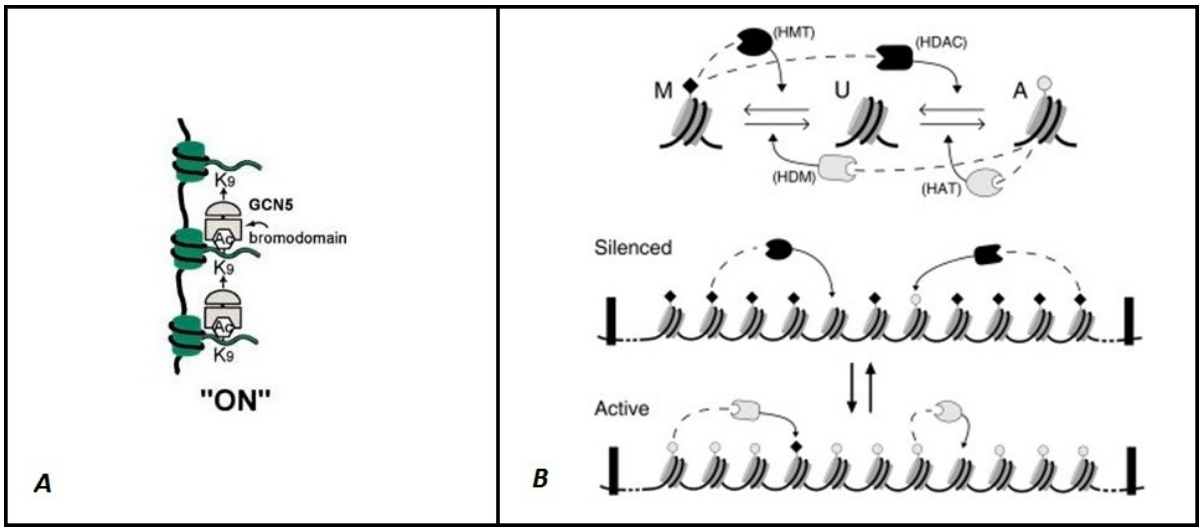


Figure 7 **Feedback loops in chromatin.** (A) Example of positive feedback in chromatin signaling mediated by a bromodomain [60]. (B) Model mechanism proposed for epigenetic memory based on positive feedback loops in nucleosome modification. Black diamond, methylation; gray circle, acetylation; HMT, histone methyltransferase; HAT, histone acetyltransferase; HDM, histone demethylase; HDAC, histone deacetylase [63].

phosphorylation events create docking sites for effector molecules, which in turn propagate and/or maintain signals through the establishment of feedback loops [61].

An example of a signaling cascade and establishment of a feedback loop in chromatin modification is shown in **Figure 7A**. Gcn5 is a bromodomain-containing HAT. The bromodomain is a PTM recognition motif, which specifically binds to acetyl-lysines. After Gcn5-dependent acetylation, other Gcn5 molecules can be targeted by the Gcn5 bromodomain to the newly acetylated lysines. Subsequent acetylation-dependent recruitment of Gcn5 molecules in the same region of chromatin can lead to propagation of acetylation marks and establishment of an activated state of chromatin (positive feedback loop) [60].

Another important aspect in chromatin post-translational modification is the epigenetic memory, which is defined as the inheritance of histone modification states through cellular generations. One can think for example of the pattern of PTMs associated with the hereditable inactivation of one of the two chromosomes X in mammals [62, 63].

A proposed mechanism for epigenetic memory is based on the establishment of positive feedback loops in nucleosome modification, similar to the ones postulated in the signaling network hypothesis **(Figure 7B)**. It is proposed that nucleosomes that carry a particular modification can become docking sites for enzymes which catalyze a similar modification on neighbouring nucleosomes. This leads to the stable recruitment of a specific set of enzymes to defined chromatin regions. Clusters of nucleosomes are then stably

mantained in a particular modification state over time. These states are proposed to be inherited through DNA replication and hence to propagate throughout generations [63].

1.3.3. Protein modules that recognize PTMs

There are two proposed mechanisms by which a cell can sense specific PTMs and convert them into *ad hoc* downstream events: the "direct" and the "effector-mediated" model [64]. On one hand, according to the "direct" model, PTMs directly affect chromatin compaction by introducing changes in the charge of histone residues or by adding steric constraints. On the other hand, the "effector-mediated" model proposes that histone PTMs can be "read" by protein modules called effectors, which in turn facilitate specific downstream events via recruitment or stabilization of module-associated proteins and complexes.

Acetylation, methylation and phosphorylation are the best studied histone modifications. The effector modules which mediate their recognition are very well-known.

1.3.3.1. Bromodomains

The bromodomain (BRD) comprises a region of ~110 amino acids, which is conserved in eukaryotes and is commonly found in transcriptional regulatory proteins. This domain is usually found once per protein, but sometimes it can be present in repeats. BRDs are often present in proteins that contain other conserved domains, especially in HATs and in ATP-dependent remodeling enzymes [65]. There are about 42 BRD-containing proteins in the human genome, with a total of 56 unique BRDs. The overall sequence similarity between members of the BRD family is not high and 9 groups of BRDs can be distinguished within the family (each group shares similar sequence length and at least 35% sequence identity) [66].

The role of BRDs was first discovered in 1999 with the NMR solution structure of the BRD of the HAT co-activator PCAF (p300/CBP associated factor) [67]. In this paper, Dhalluin *et al.* showed by using NMR that the PCAF BRD acts as an acetyl-lysine (AcLys) binding module. Several structures of different BRDs in complex with mono-acetylated histone peptides have been reported to date. Interestingly, a single BRD binding two acetylation marks was described as well [68]. Despite the overall low sequence similarity among BRD modules, there is a high degree of conservation in the general bromodomain

structural fold and in the residues responsible for the acetyl-lysine recognition **(Figure 8)** [69].

BRDs adopt a conserved structural fold of a left-handed four-helix bundle (helices α_Z, α_A, α_B, α_C), with two long inter-helical loops of variable length and sequence (termed ZA and BC loops). The loop ZA between the helices α_Z and α_A packs against the loop BC between the helices α_B and α_C and forms a hydrophobic pocket, which stabilizes the BRD structure and serves as a binding pocket for the acetyl-lysine [69]. In general, sequence similarity in this region is very low, with the ZA and BC loop being the region of major sequence variation within BRDs. However, the residues engaged in acetyl-lysine recognition are highly conserved. In particular, all BRDs have a highly conserved asparagine residue in the binding site (Asn803 in PCAF, Asn408 in GCN5, Asn1168 in CBP and Asn1131 in p300) which mediates a hydrogen bond with the N-acetyl carbonyl oxygen of the acetyl-lysine [70]. Binding specificity among BRDs is thought to be achieved by sequence variations in the ZA and BC loops. These loops can determine specificity of binding by interacting with residues flanking the acetylated lysine (Kac -/+1,-/+2,-/+3) [70].

The hydrophobic pocket, which is the acetyl-lysine binding site, is located at one end of the four-helix bundle, opposite to the N- and C- termini of the domain. This explains how different BRDs can be found sequentially within a protein (tandem or multiple BRDs) and can function as distinct structural units for protein-protein interactions [67, 69]. There are several roles postulated for BRD-AcLys interactions. First, BRDs may stabilize HATs on acetylated chromatin, leading to an increase in local levels of acetylation. Second, they

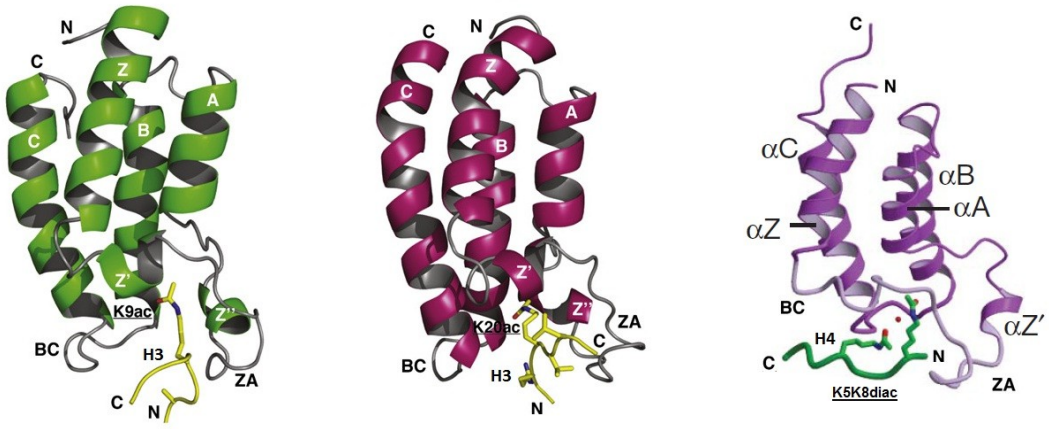


Figure 8 **Structure of bromodomains bound to acetylated histones.** The overall structure of the bromodomains is conserved. From left to right, the bromodomains of PCAF in complex with H3K9ac, of CBP in complex with H3K20ac and of Brdt-BD1 in complex with H4K5K8diac. It is interesting to note that Brdt-BD1 can bind two acetylation marks simultaneously. The loop ZA an BC are indicated. Modified peptides are shown in yellow (PCAF and CBP) and green (Brdt-BD1) sticks [68, 70].

could target BRD-containing complexes to chromatin or to non-histone proteins. Third, they could act in protein autoregulation, by binding intramolecularly to autoacetylated HATs. Last, they may act in combination with other effector modules to mediate specific downstream outcomes [71, 72].

1.3.3.2. PHD fingers

The PHD finger (plant homeodomain) has been established as a *bona fide* domain in the early nineties when it was found to be present in a wide variety of eukaryotic proteins [73]. It comprises a cysteine-rich region of ~60 amino acids with a unique Cys_4-His-Cys_3 (four cysteines, one histidine, three cysteines) pattern. These residues coordinate two Zn^{2+} ions in a cross-brace fashion. In general, the PHD finger adopts a globular fold, consisting of a double-stranded anti-parallel β-sheet and two small C-terminal α-helices. There are two loops between the structured regions (loop 1 and 2) which are variable in sequence and thought to contribute to specificity of PHD fingers **(Figure 9)** [10].

The PHD finger is different but closely related to the RING domain (Really Interesting New Gene), which is characterized by a Cys_3-His-Cys_4 pattern and also binds two Zn^{2+} ions. RING domains are known to mediate the ubiquitination of proteins and it has recently been reported that PHD fingers may act in a similar way to promote protein sumoylation [74]. However, the majority of PHD finger-containing proteins in eukaryotic genomes are found to be nuclear and predicted to bind chromatin or nuclear proteins [10]. In recent years, a role of PHD fingers in mediating protein-protein interactions through recognition of specific PTMs has become evident. PHD fingers can be divided in different subsets based on their substrate binding specificity **(Figure 9)** [75].

The first identified PHD subset comprises PHD fingers capable of binding a tri-methylated lysine in position 4 of histone H3 (H3K4me3). These PHDs are highly specialized and can distinguish among different degrees of methylation (with a decreasing affinity which goes from strong binding for Kme3 to no binding for Kme1 and Kme0). This specificity is due to the peculiar amino acid composition in the PHD binding pocket. The recognition of H3K4me3 involves two distinct surface channels for histone binding: the K4me3 and the R2 (arginine in position 2) binding pockets. The trimethyl-ammonium group of K4me3 binds into an aromatic cage and a conserved tryptophan separates the K4me3 and the R2 binding pockets. A network of hydrogen bonds and complementary surface interactions complete the interface where the histone peptide adopts an extended

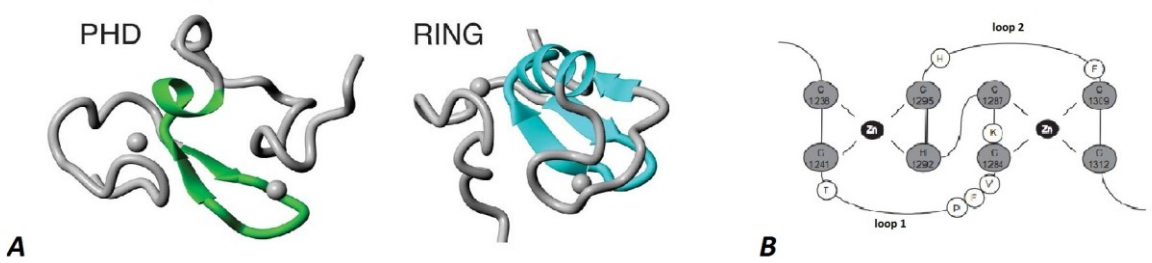


Figure 9 **Structural comparison of zinc finger folds.** (A) The PHD finger motif from William syndrome transcription factor (PDB: ID 1F62; left) and the RING domain from equine herpes virus-1 (PDB: ID 1CHC, right) are shown [112]. (B) A schematic representation of the CBP PHD finger is shown. The residues coordinating the zinc-atoms are represented in grey circles [117].

conformation and contributes another strand to the anti-parallel β-sheet [64]. This arrangement requires a separation of only one amino acid between the arginine and the lysine, which is unique for the ARTKme3QT sequence in the histone H3 tail. Binding of H3K4me3 by PHD fingers can lead to very different downstream events. According to the specific cellular context, it can mediate transcriptional activation, repression or recombination [59, 76, 77].

A second major subset of PHD fingers can bind the unmodified histone H3 tail (unH3, ARTK sequence). In this case, the unmodified histone peptide adopts the same extended conformation as the trimethylated one in the H3K4me3 case. The unmodified peptide is recognized and binds in a hydrogen bond cage formed by three backbone carbonyls. The H3 Lys4 forms two additional hydrogen bonds, one of the two with an aspartate side chain. The recognition specificity is achieved by steric hindrance, since the binding pocket would not be able to accommodate extra methyl groups on lysines. As in the case of H3K4me3, PHD recognition of unmodified lysines can lead to very distinct outcomes [64, 75, 78].

Other subsets of PHD fingers are less well characterized. They can be involved in recognition of trimethylation on lysine 9 of histone H3 (H3K9me3), a mark usually enriched at transcriptionally inactive genes, or they can recognize acetyl-lysines on histone H3 and H4 [75, 79].

1.3.3.3. Other effector modules

Two classes of protein modules are known to recognize methyl-marks and share several recognition features, such as the presence of an aromatic cage for binding with the methyl-ammonium group. These are the PHD fingers (discussed in the previous section) and the members of the Royal superfamily of folds. For both classes, hydrogen bonding and steric hindrance are increasingly important to discriminate between lower methylation states [80].

Among the Royal superfamily, the first documented effector module has been the chromodomain. Chromodomains appear not to be selective between degrees of methylation, but are usually not able to bind unmethylated-lysines [64]. Some chromodomains do not bind to histones but instead can target RNA and DNA [81]. Other members of the Royal superfamily are the Tudor domains. Tudor domains are known to recognize methyl-arginine residues and to discriminate between symmetric and asymmetric methylation. Double-Tudor domains have been shown to bind to methylated lysines and they are able to select among degrees of methylation [80]. MBT (Malignant Brain Tumour) domains are also part of the Royal superfamily. They consist in repeats of ~70 aminoacids and they are known to target mono- and di-methyllysines [64].

A third class of methyl-lysine binding modules is represented by the WD40 repeats. These repeats are characterized by a seven bladed β-propeller with a channel whose surface constitutes the binding site for the histone peptide. The specificity of H3K4 binding is achieved by an extensive interaction network between the WD40 surface and several peptide side chains as well as the N-terminal amino group. There is no aromatic cage for methyl-lysine binding. WD40 repeats are usually associated with H3K4me2 recognition *in vivo* [80].

Effector modules responsible for the recognition of phosphorylated residues are also important. There are very well-known phosphate-binding modules described for non-histone proteins (such as the SH_2 domains in protein kinases). However, little is known about modules which recognize histone phospho-marks. Among the few examples reported to date, the 14-3-3 proteins constitute a family of conserved phosphoserine binding modules, which recognize phosphorylated serine 10 on histone H3 (H3S10pho) [64].

Finally, the SANT domain is a relatively small domain of about 50 aminoacids. It has been reported as a putative DNA-binding domain. However, it can be found in many ATP-dependent remodelling enzymes and other chromatin-related proteins. SANT domains can bind directly and stabilize the histone N-terminal tail, facilitating the binding of SANT-

associated proteins to chromatin. These effector modules interact preferentially with unmodified histone tails [82].

1.3.3.4. Linked binding modules

There are several examples of effector modules linked within the same protein or the same multisubunit complex. These linked modules may act combinatorially to allow recognition of multiple PTMs or display new functions that are not possessed by the single modules. Multivalency, that is inherent in such systems, likely results in increased affinity for substrates and high specificity [64, 83].

There are several examples of structurally and functionally characterized paired modules. In human proteins, BRDs are often found in combination with other BRDs or with PHD fingers. Also, PHD-PHD combinations are possible [66].

Linked modules can sometimes serve to read multiple marks residing on the same or a different molecule. The double BRD of TAF1 (TBP associated factor 1), for example, can bind multiple acetyl-lysine marks simultaneously. It also shows increased affinity for di-/tetra-acetylated substrates as compared with mono-acetylated ones [84]. Another interesting example is the PHD-BRD module of BPTF (bromodomain and PHD domain transcription

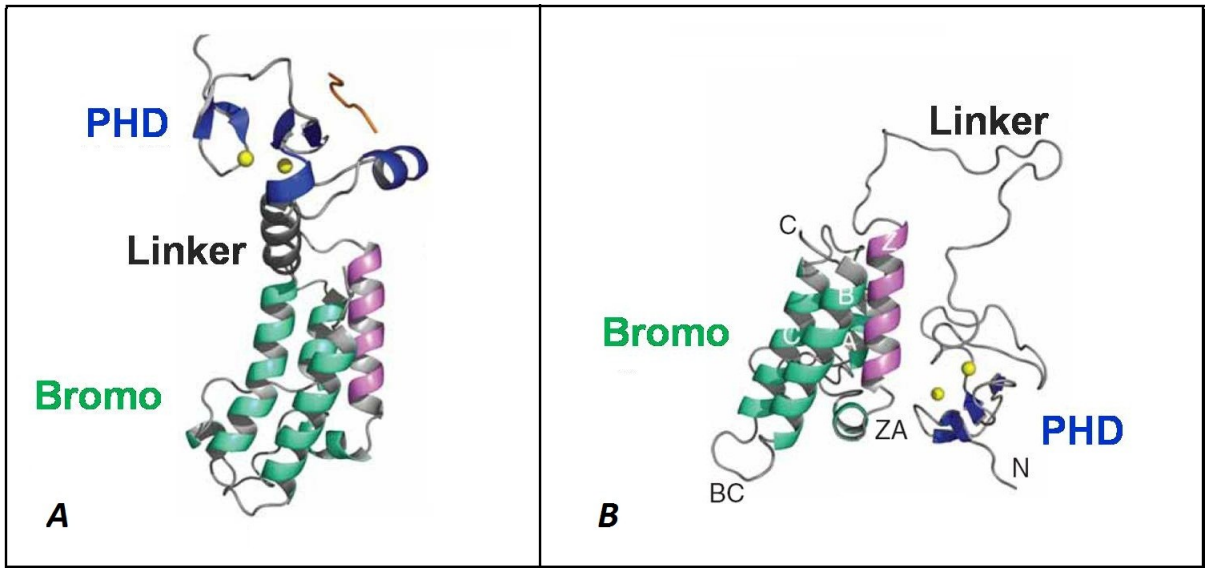


Figure 10 **Structures of PHD-bromo modules.** (A) Structure of PHD-BRD of BPTF. The two effector modules are separated by a rigid helical linker and can function independently in substrate targeting. (B) Structure of the PHD-BRD of KAP-1. Here, the α_Z helix forms a central hydrophobic core which anchors the PHD domain on one side and the other three helices of the BRD on the other side. The PHD finger and the BRD in KAP1 cooperate as one functional unit to facilitate lysine sumoylation [87].

factor). This module has the potential to recognize K4me3 on histone H3 (PHD domain) and to combinatorially bind to acetylated lysines on histone H4 (BRD) **(Figure 10A)** [76].

Moreover, linked modules can serve to create unique binding site, such as in the case of the double PHD in DPF3 (D4, zinc and double PHD fingers, family 3). The two binding surfaces of the DPF3 PHD fingers are combined to form an extended surface groove, creating a single binding site for acetylated lysines [85].

Finally, linked modules can be involved in protein autoregulation (double bromodomain of Rsc4) [86] or they may function independently of PTM recognition. The KAP1 (KRAB associated protein 1) PHD-BRD module is a good example of the latter case. This module functions as an E3 ligase and directs intramolecular sumoylation within the KAP1 BRD. This modification results in transcriptional silencing **(Figure 10B)** [87].

It is important to note that similar combinations of effector modules (*i.e.* double BRDs, PHD-BRD modules) can lead to very distinct outcomes. It has been shown for the PHD-BRD modules of BPTF and KAP1 that the different roles played by these effector modules in the two proteins are also reflected in significant differences in their 3D structure **(Figure 10)**. A similar module (BRD-PHD) is also present in the p300/CBP acetyltransferase. Structural characterization of this module is not available to date. However, there are studies showing that bromo- and PHD-domains in p300/CBP can cooperate to robustly associate with hyperacetylated chromatin [88].

1.4. ACETYLTRANSFERASES AND DEACETYLASES

Among the enzymes that carry out post-translational modification of histones, the histone acetyltransferases (HATs) and the histone deacetylases (HDACs) are the best studied ones. Acetyltransferases are usually associated with transcriptional activation, whereas deacetylases promote gene silencing. The balance between acetylation and deacetylation of histone and non-histone proteins controls gene expression and a variety of cellular processes. Therefore, the aberrant activity of acetyltransferases and deacetylases has often been implicated in several human diseases, such as metabolic disorders and cancer, thus making them important drug targets [51].

1.4.1. Acetyltransferases

Acetyltransferases mediate the transfer of acetyl-groups onto the ε-amino group of lysine side chains [51]. Many transcriptional co-activators are known to have HAT activity and to be present *in vivo* within multiprotein complexes [89]. Interestingly, many HATs have transcription factor acetyltransferase activity (FAT), as they can catalyze the acetylation of non-histone proteins.

There are three major families of acetyltransferases: the GNAT family (Gcn5-related N-acetyltransferases), the MYST family (named from its founding members: MOZ, Ybf/Sas3, Sas2, Tip60) and the p300/CBP family (protein of 300 kDa and CREB binding protein) [90]. There is high sequence similarity within these families but very poor similarity among them. Moreover, each family has generally unique substrate specificity and appears to mediate different biological functions. However, all known HATs contain a catalytic domain (called HAT domain) which varies in size between families and appears in association with different sets of effector modules **(Figure 11)** [89].

For example, the GNAT family contains a 160 residues HAT domain and a highly conserved bromodomain. In contrast, the MYST family contains a 250 aa HAT domain including a zinc-binding domain and an N-terminal chromodomain. Finally, the p300/CBP family comprises proteins which are more global regulators of transcription and contain several effector modules (including a BRD and four zinc-finger motifs). The HAT domain of this family is larger and comprises a region of about 500 aminoacids [89].

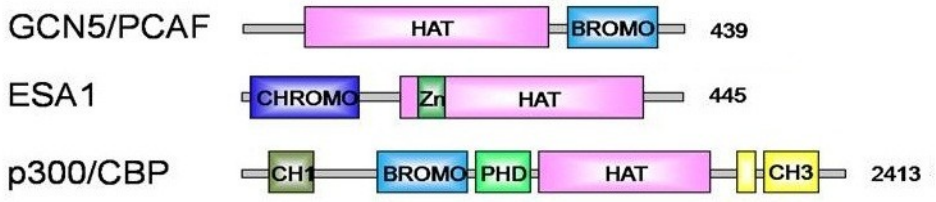


Figure 11 **Cartoon showing organization of different HAT families.** Members of the GNAT family (GCN5/PCAF), the MYST family (ESA1) and the p300/CBP family (p300/CBP) are shown. The presence of effector modules neighbouring the catalytic site is common in all HATs.

1.4.1.1. Catalytic mechanism

In all the families of HATs, there is a conserved central core domain which mediates not only AcCoA (acetyl coenzyme A) binding but also catalysis **(Figure 13A)**. The GNAT family uses a ternary complex mechanism of acetyl-transfer involving a conserved Glutamate residue (E173 in Gcn5) within the core domain. In this mechanism, AcCoA and the histone substrate bind the enzyme and form a ternary complex before catalysis. After binding, the conserved glutamate residue in the core domain assists the deprotonation of the ε-amino group of the target lysine and allows nucleophilic attack on the carbonyl carbon of AcCoA. Finally, the two reaction products, CoA and acetylated protein, can be released **(Figure 12)** [51].

For long time it was reported that the MYST family of HAT was using a different catalytic mechanism for acetyl-transfer, involving the formation of an acetyl-enzyme (acetyl-cysteine) intermediate: the ping-pong mechanism [51]. Recently it was shown that the MYST family can catalyze efficient acetyl-transfer without the need of this acetyl-cysteine intermediate. Instead, a conserved glutamate residue (E338) functions in a similar way as the E173 in Gcn5, suggesting that the catalytic mechanism is conserved [91].

The p300/CBP family as well does not require the formation of an acetyl-enzyme intermediate for catalysis. However, this family of HATs uses a specific type of sequential catalytic mechanism, which involves the formation of a non stable ternary complex: the Theorell-Chance (or 'hit-and-run') mechanism **(Figure 12)**. Following AcCoA binding, the peptide substrate associates only transiently with the enzyme to allow the acetylation

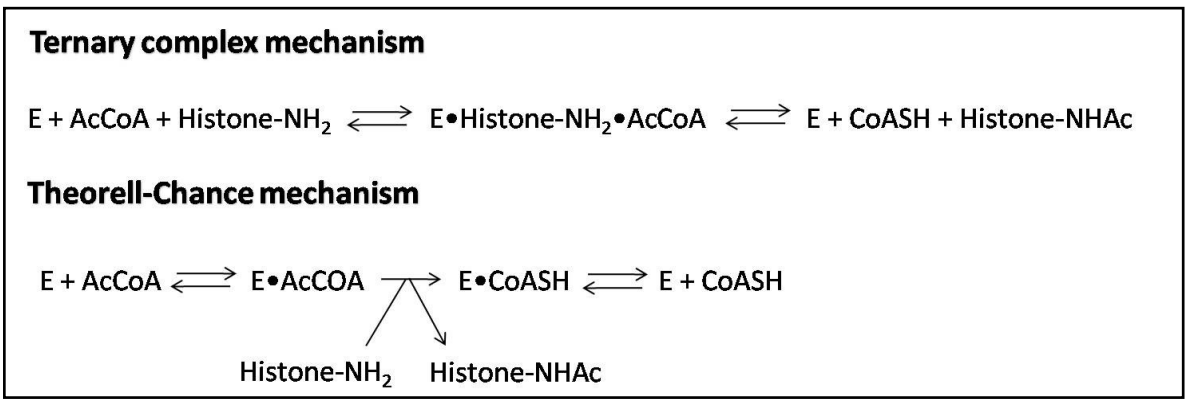


Figure 12 **Catalytic mechanisms of HATs.** HATs do not require the formation of an acetyl-enzyme intermediate for catalysis. HATs from the GNAT and MYST family catalyze the transfer of the acetyl group through a ternary complex mechanism. HATs from the p300/CBP family use a different kind of mechanism often referred to as Theorell-Chance mechanism (or 'hit and run') which requires the formation of a non stable ternary complex. "E" stays for enzyme.

reaction and departs immediately after the acetyl-transfer. The p300/CBP family does not depend on a specific glutamate residue as a base for catalysis, unlike the GNAT and MYST families. Instead two other residues are important for activity: a tyrosine residue (Tyr1467 in p300), which acts as a general acid to protonate the CoA leaving group, and a tryptophan residue (Trp1436 in p300), which helps the correct positioning of the lysine into the catalytic pocket [92].

1.4.1.2. Overall structure of the HAT domain

Structures of HAT domains have been firstly described for the members of the GNAT family [51, 93]. GNAT HAT domains contain three sequence motifs (called A, B and C) [94]. These motifs form a structurally conserved central domain composed of three antiparallel β-strands followed by an α-helix and an additional β-strand region **(Figure 13A)**. This domain mediates the interaction with the acetyl-coenzyme A (AcCoA) cofactor. The flanking C- and N- terminal regions play an important role in histone substrate binding [89].

The structures of HAT domains from several members of the other HAT families have also been determined to date [51, 92, 95]. A comparison of these structures reveals that the core domain found in the GNAT family (A,B,C motifs) is conserved between HAT families, suggesting a similar mechanism for the recognition of the AcCoA and for substrate catalysis **(Figure 13A)**. In contrast, protein segments located N- and C- terminal of this conserved core domain show structural divergence. It is proposed that these domains are involved in substrate binding [89].

The p300/CBP HAT domain structure is described in part 1.5. It is important to point out that there are two more features that distinguish the p300/CBP family from the other HATs. Firstly, the surface charge of the p300/CBP HAT domain is highly negative, particularly in proximity of the catalytic site. The other HATs are characterized by more neutral surface charge in this region **(Figure 13B-D)**. Secondly, the AcCoA binding site is more buried than in other HATs due to the presence of a specific loop (L1) which covers part of the cofactor [92].

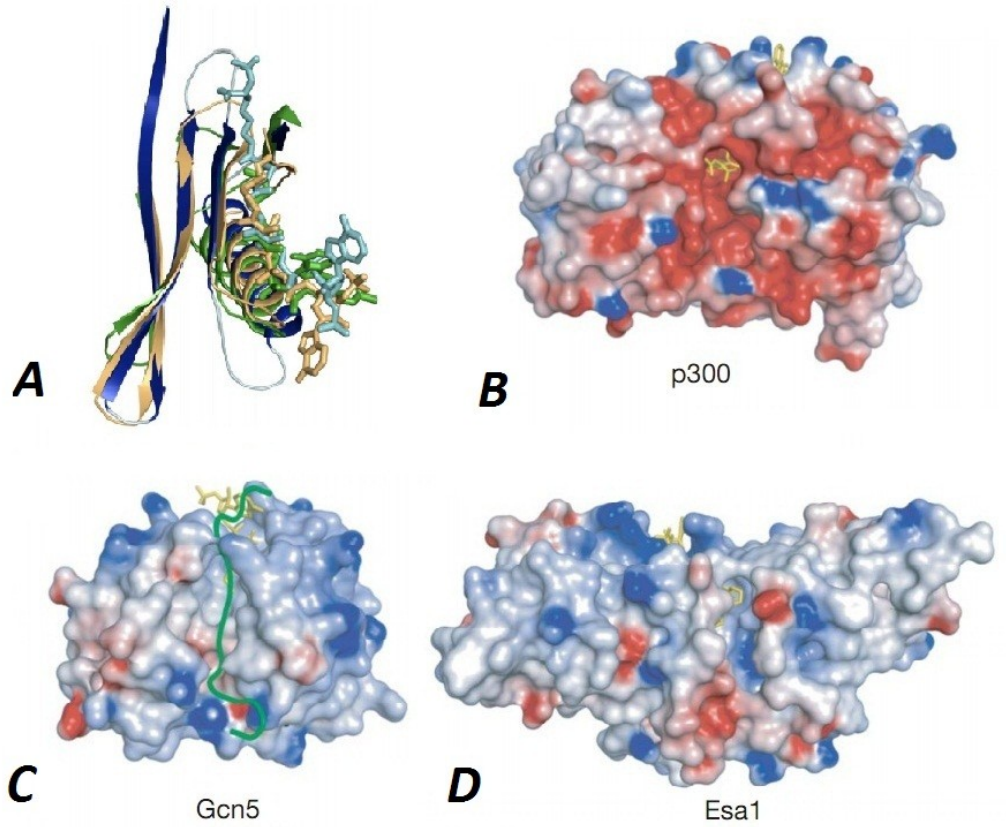


Figure 13 **Structure of the HAT domain.** (A) Superposition of the core domain of three HATs (A,B,C motifs): p300 (blue), Gcn5 (green), Esa1 (orange). (B-D) Electrostatic surface representation of the HAT domains of p300, Gcn5 and Esa1 respectively. LysCoA in p300 and CoA in Gcn5 and Esa1 are represented with yellow sticks. The histone substrate of Gcn5 is shown in green [92].

1.4.2. Deacetylases

Deacetylases mediate the removal of acetyl-groups from the ε-amino group of lysine side chains [51]. Based on sequence similarity and cofactor dependency, HDACs have been subdivided in four different classes and two different families. The classical HDAC family is composed of the class I HDACs (comprising HDAC1, -2, -3, -8), the class II (including HDAC4, -5, -6, -7, -9, -10) and the class IV (related to human HDAC11). Class I, II and IV share sequence similarity within their catalytic domain and require a Zn^{2+} ion as a cofactor for the enzymatic activity. The sirtuin family contains members of the class III HDACs which are yeast Sir2 (Silent information regulator-2) homologues. They do not share

sequence homology with the members of the classical HDAC family and use nicotinamide adenine dinucleotide (NAD+) as a cofactor for catalysis [51, 96].

A aforementioned, the balance between acetylation and deacetylation of histone and non-histone proteins controls gene expression and a variety of cellular processes. The interplay between sirtuins and p300/CBP for reciprocal activity regulation have been extensively documented [49, 97-99].

1.4.2.1. The sirtuin family

Homologues of yeast Sir2 were found in the three domains of life bacteria, archea and eukaryotes and show deep conservation. While bacteria and archea contain one or two sirtuins, eukaryotes contain multiple members. For instance, in human there are seven Sir2 homologues called SIRT1-7 [100]. Sirtuins are associated with histone H4 hypoacetylation at lysine 16, which is a hallmark of cancer [57]. However, their substrate specificity extends beyond histones and allows for participation in several biological processes. In mammals, sirtuin activity has been linked to counteracting age-associated disorders, such as neurodegenerative diseases, obesity and type II diabetes [96].

Sir2-activating compounds are currently being used as chemotherapeutic agents and there is increasing therapeutic interest in the sirtuin family [51]. To this purpose, many efforts were done in determining the mechanism of catalysis and the structure of Sir2 proteins.

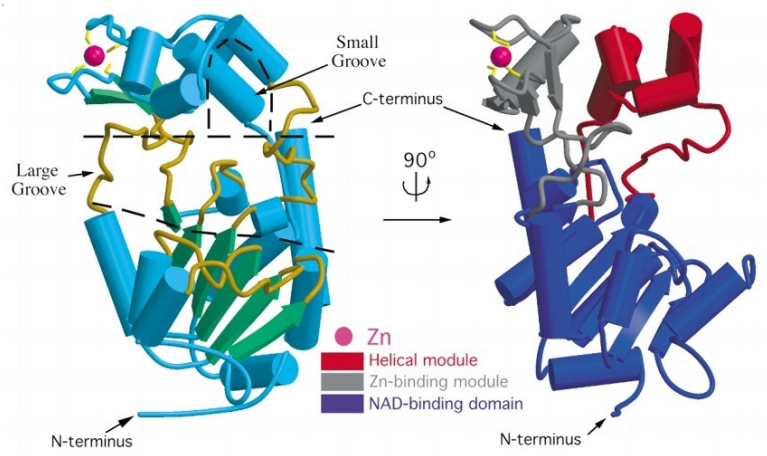


Figure 14 **Structure of the SIRT2 deacetylase.** Two views of the overall structure of the catalytic core of SIRT2 rotated by 90°. The loops connecting larger and smaller domain are represented in yellow on the left. On the right, the Rossman fold is in blue [101].

The overall structure of sirtuin proteins reveals a highly conserved catalytic core of about 270 amino acids. The C- and N-terminal regions of this core are highly variable between the family members and are thought to be important for protein specificity. The sirtuin catalytic core has an elongated shape and contains two domains: a larger domain that is a variant of the Rossman-fold domain, a common feature of NAD(H)/NADP(H)-binding proteins, and a smaller domain that contains a structural zinc atom and a hydrophobic pocket. A series of loops between the two domains forms a large groove which includes the NAD+ binding site and the catalytic site of the protein. Within the smaller domain there is a small groove with conserved hydrophobic residues, which is contiguous to the large groove and located close to the NAD+ binding site. This small pocket is considered to be the putative substrate binding site of the protein **(Figure 14)** [100, 101].

The precise chemical mechanism of deacetylation catalyzed by Sir2 proteins is still under debate. However, some aspects of the reaction are generally accepted. Deacetylation by sirtuins needs acetyl-lysine substrate and NAD+ cofactor for catalysis. The reaction mainly involves three coupled steps. Firstly, the enzyme binds the NAD+ cofactor and an acetylated lysine substrate. Secondly, the sirtuin catabolises the NAD+ molecule into activated ADP ribose and nicotinamide. An enzyme-ADPribose-substrateAcLys intermediate is formed. The nicotinamide (which is as well a non competitive inhibitor of the sirtuin reaction) is released. Thirdly, the sirtuins catalyse the transfer of the acetyl-group to the ADP-ribose. The final products of the reaction are the deacetylated substrate and *O*-acetyl-ADP ribose. The latter compound can also act downstream the deacetylation reaction as a signalling molecule **(Figure 15)** [57].

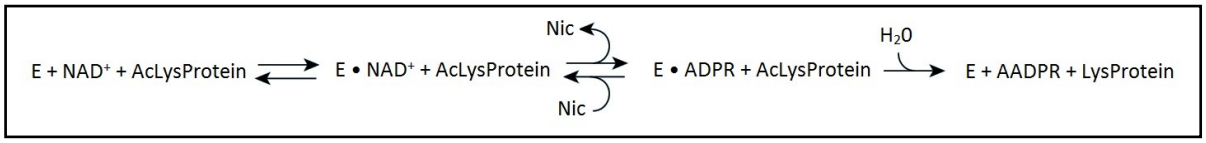


Figure 15 **Sirtuin enzymatic reaction.** The deacetylation reaction is composed of three sequential steps. The first two steps are reversible, whereas the final step is irreversible. Details about the reaction are discussed in the text. E, enzyme; ADPR, ADP-ribose; AADPR, O-acetyl-ADP-ribose; Nic, nicotinamide.

1.4.2.2. The SIRT2 deacetylase

The mammalian members of the Sir2 family (SIRT 1 to 7) comprise a very heterogeneous group of proteins that are involved in different functions and localize in several distinct compartments within the cell. Indeed, SIRT4 and -6 do not seem to be active as deacetylases and rather only exhibit mono-ADP-ribosylation activity. In contrast, SIRT1 to -3 exhibit HDAC activity and are specific for important residues such as H4K16ac, H3K9ac or H3K56ac. Moreover, mammalian sirtuins can be cytoplasmic, nuclear or mitochondrial [57].

SIRT2 is mainly cytoplasmic. However, it was shown that SIRT2 can undergo constant shuttling inside and outside of the nucleus and that it is predominantly nuclear during mitosis [57, 102]. This is indeed important so that SIRT2 can carry out the deacetylation of histones at specific positions, such as H4K16ac and H3K56ac. H4K16ac and H3K56ac are implicated in gene transcription regulation (together with chromatin decondensation and DNA repair, respectively) and can be catalyzed by several HATs, including the p300/CBP acetyltransferase. The opposite enzymatic activities of p300 and SIRT2 suggest that these enzymes may interplay in the regulation of specific genes [49, 57].

As discussed above, the balance between acetylation and deacetylation controls gene expression. This equilibrium implies that acetyltransferases and deacetylases are implicated in similar regulatory phases and can affect each other. In the case of SIRT2 and p300, a direct cross-talk has been documented in addition to their ability to target the same histone residues [98, 99].

SIRT2 can bind p300 in the region from the BRD to the CH3 domain. Upon interaction, SIRT2 and p300 can affect each other directly in two major ways. On one hand, p300 can acetylate SIRT2 and down-regulate its deacetylase activity. This can have important consequences in transcriptional regulation, given the ability of SIRT2 to deacetylate p53 and repress its activity [99]. On the other hand, SIRT2 can deacetylate p300 within its autoinhibitory loop region and have an impact on pre-initiation complex (PIC) assembly. As a consequence, p300 can undergo cycles of autoacetylation/deacetylation assisted by SIRT2 [98].

1.4.3. Acetylation vs phosphorylation

There are several aspects which are common to both acetylation and phosphorylation events. It is known that phosphorylation can regulate key cellular processes in response to cellular signals. Similarly, acetylation can also play an important role in signal transduction. Both modifications are ubiquitous and affect nuclear and cytoplasmic proteins. They also regulate diverse functions (such as DNA recognition and protein-protein interactions) and some cellular processes (such as cell-cycle and DNA transcription) [1, 3, 103].

The analogy between phosphorylation and acetylation can go forward, considering the signalling aspect of these modifications. Although a cascade implying solely acetylation events has not been reported yet, there are examples of acetylation cascades which can combine with phosphorylation cascades [103, 104]. Moreover, acetyltransferases and

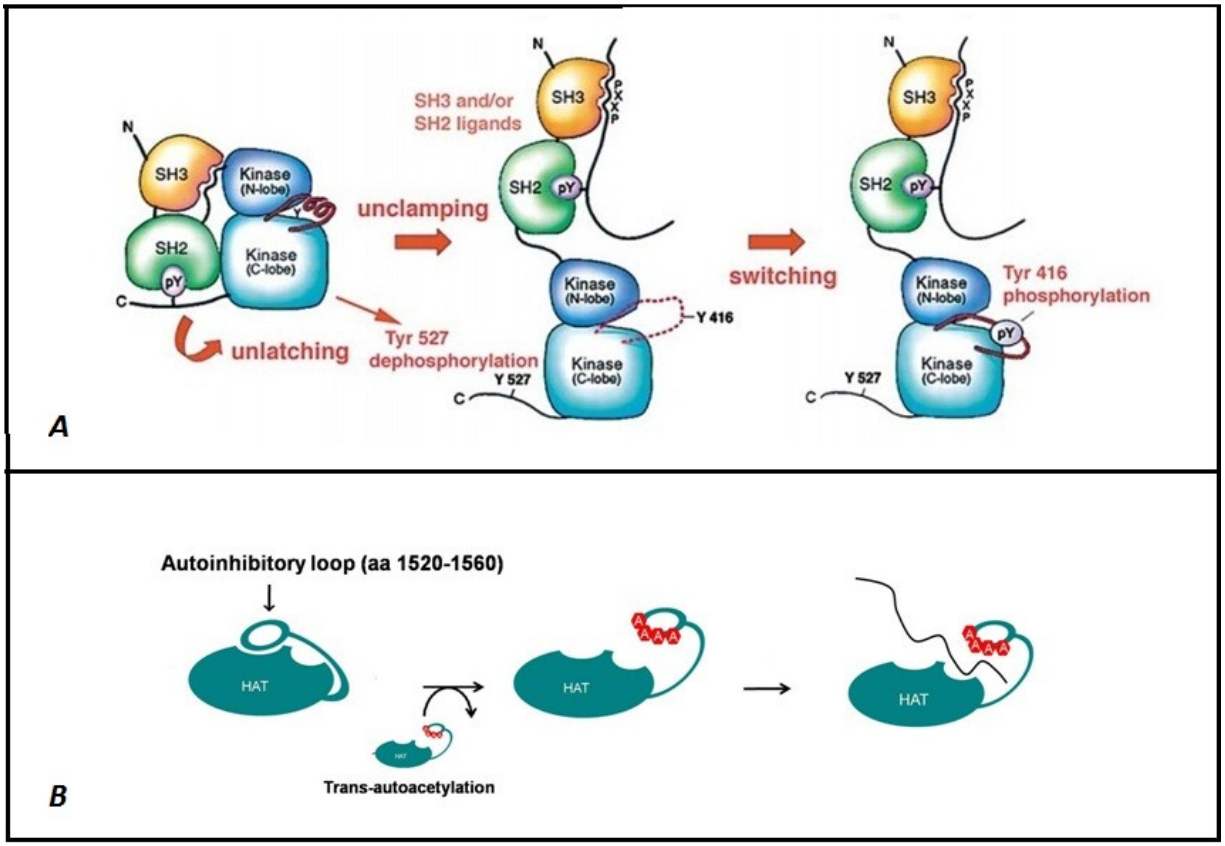


Figure 16 **Kinases vs acetyltransferases.** (A) Mode of activation of Src family kinases. The assembled state is unlatched by dissociation of the C-ter from the SH2 domain. Competing SH2 and SH3 substrate can unclamp the assembled state. Phosphorylation in the inhibitory loop switches the kinase domain into its active form [8]. (B) Mode of activation of the p300/CBP acetyltransferase. Trans-autoacetylation within the autoinhibitory loop stimulates HAT activity.

kinases contain specific sets of effector modules and can act as multivalent adaptor proteins, similarly to many signalling proteins. Finally, both acetylated and phosphorylated residues can create docking sites for signalling proteins [105]. The BRDs of acetyltransferases, which recognize acetyl-lysines, may be analogous to SH2 and SH3 domains in protein kinases, which bind phospho-tyrosines and mediate the transmission of the phosphorylation signal [103].

Protein kinases are usually tightly regulated by inhibitory loops which can limit accessibility of the substrate to the binding site. Accessibility can be regulated by phosphorylation and dephosphorylation events occurring within these loops **(Figure 16A)**. Activation can occur through autophosphorylation (in *cis* or *trans*) or can be directed by another kinase in the regulatory cascade. Moreover, effector modules present in kinases can act *in cis* and mediate phosphorylation-dependent autoregulation of the enzyme activity [8]. Similarly, some acetyltransferases can utilize an inhibitory loop to regulate substrate accessibility to the active site. For example, an autoinhibitory loop is present within the HAT domain of p300/CBP and can activate the enzyme upon autoacetylation **(Figure 16B)** [6]. Moreover, it was reported that rsc4 (a subunit of the RSC complex, Remodels the Structure of Chromatin) is regulated in an acetylation- dependent manner and a double-BRD was shown to be involved [86].

1.5. THE p300/CBP ACETYLTRANSFERASE

Two members of the p300/CBP family of acetyltransferases are the human protein p300 (2414 aa) and its homologue CBP (2441 aa). CBP and p300 are usually referred to as the protein pair p300/CBP because they share 91% sequence identity and are thought to be functionally equivalent [106].

The p300/CBP acetyltransferase is an important transcriptional coactivator which is involved in a wide range of biological processes including DNA transcription, development, innate immune response and cell cycle regulation [3, 4]. Due to the versatility and the broad specificity of p300/CBP, the cellular demand for this acetyltransferase can be so high that the coactivator may be used competitively by different pathways and represents a limiting factor (and a regulator) for transcriptional activity [107].

There are three major mechanisms of transcriptional activation described for p300/CBP. Firstly, it can activate transcription through acetylation of histone or non-histone proteins. Secondly, it can act as a multivalent scaffold to recruit other cofactors or to allow

the assembly of multiprotein complexes. Thirdly, it can serve as a bridge to connect sequence-specific transcription factors to the components of the basal transcriptional machinery **(Figure 17A)** [3].

These functions of p300/CBP are executed through different conserved domains **(Figure 17B)**. In particular, the N- and C-terminal regions of p300/CBP act as transactivation domains and contain modules for protein interaction: a nuclear receptor-interacting domain (NR), two cysteine-histidine (CH)-rich domains (CH1 and CH3), a KIX domain (or CREB binding domain) and a glutamine/proline (QP)-rich domain comprising the IBiD domain (IRF3-binding domain). The central region of p300/CBP represents the catalytic core of the protein and contains the HAT domain and two effector modules: a BRD and a PHD (or CH2) domain [3, 107].

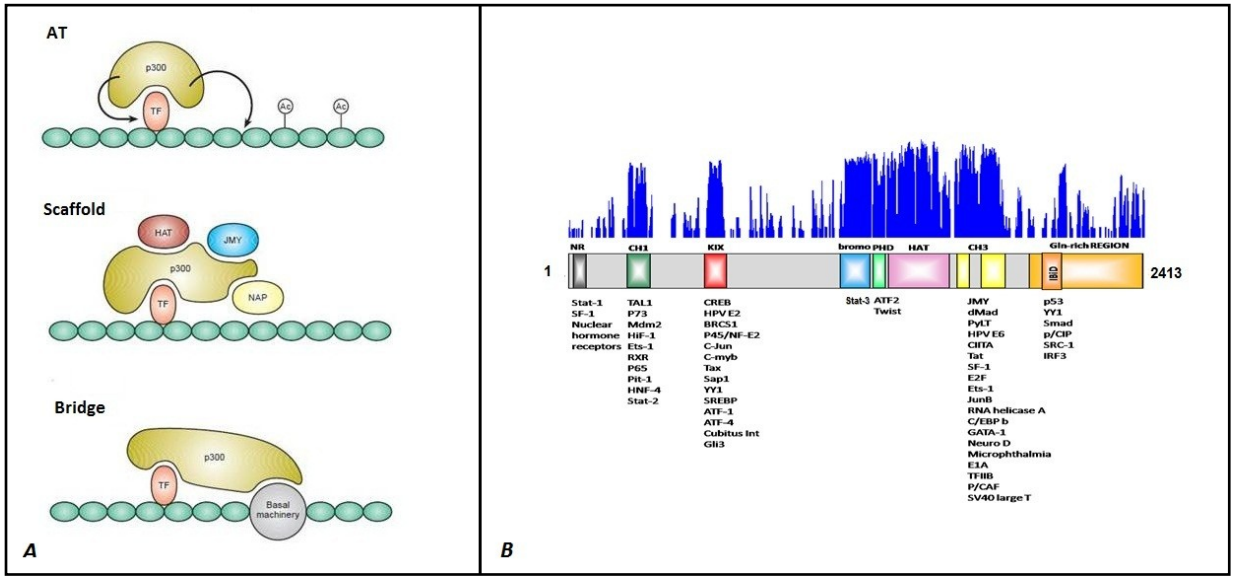


Figure 17 **The p300/CBP acetyltransferase.** (A) Three major mechanism of action of p300/CBP. Figure adapted from [3]. (B) Bar diagram of p300/CBP. Sequence conservation among higher eukaryotes is shown at the top of the bar diagram. Binding partners found for each domain are also reported.

1.5.1. The C- and N-termini of p300/CBP

The C- and N-termini of p300/CBP contain transcription activation (TA) activity. This is due to the presence of many functional domains responsible for the binding to different sets of transcriptional activators and regulators. In particular, there are at least 400 proteins that interact with p300/CBP [108]. Most of these conserved domains have been

structurally and functionally characterized [107, 108]. The N-terminal TA region of p300/CBP contains the NR (not very well characterized to date), the KIX and the CH1 domains. In the C-terminal region there are the CH3 domain and the QP-rich domain comprising the IBiD domain [107].

The KIX domain is known to bind the transactivation domain of CREB and other nuclear factors to regulate target gene expression. The structure of the KIX domain in complex with the phosphorylated transactivation domain of CREB was determined using NMR **(Figure 18A)**. The KIX domain is composed by three mutually interacting α-helices (α1, α2 and α3) and two other small α-helices (G1 and G2) connected by loops. Helices α1 and α3 form a hydrophobic patch which is the primary interacting surface of the CREB transactivation domain. The largely unfolded transactivation domain can rearrange upon binding to the KIX domain forming amphipatic α-helices and becomes active [109].

The two CH regions CH1 and CH3 comprise three zinc finger motifs: TAZ1 in CH1 and ZZ and TAZ2 in CH3. TAZ1 and TAZ2 are similar in structure and also form a structured hydrophobic core for the interaction with binding partners. Unfolded binding partners can fold upon binding to the TAZ domains in a similar way as for the KIX domain. Many cellular transcription factors and viral oncoprotein can bind to TAZ domains, which are novel zinc-binding modules. TAZ2 binds three equivalents of Zn^{2+} and folds into a compact globular structure consisting of four α-helices and three HCCC-type zinc-binding motifs. The helices pack against each other to form a hydrophobic core. Each Zn^{2+}-binding site is formed by two zinc ligands (His and Cys) located at the C-terminus of the first helix,

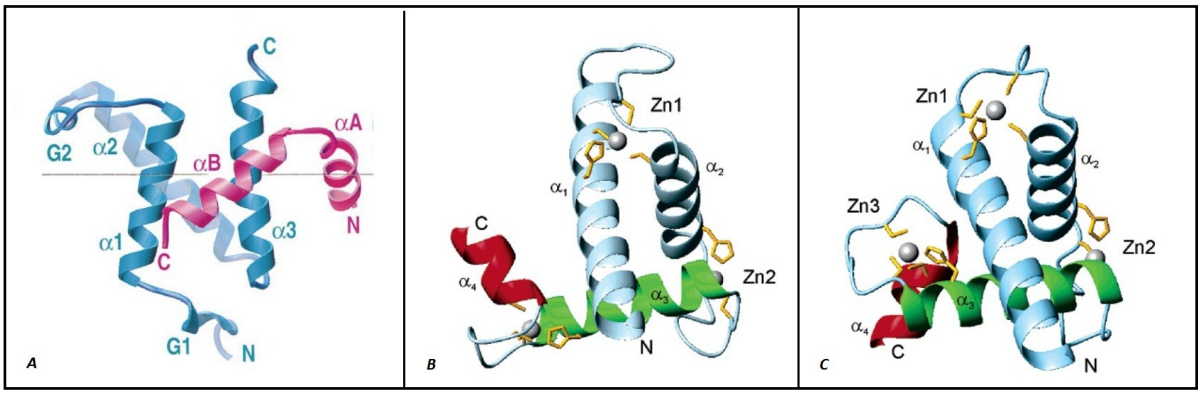


Figure 18 **Structures of p300 KIX, TAZ1 and TAZ2 domains.** (A) NMR structure of the KIX domain in complex with the transactivation domain of CREB (pKID). The backbone of KIX is shown in cyan and that of pKID in pink [109]. (B-C) NMR structures of the TAZ1 (B) and TAZ 2 (C) domains. Helices α_1 and α_2 are shown in blue, helix α_3 is in green, helix α_4 is in red. The side chains of the ligand residues in each zinc-binding site are shown in yellow [111].

the third ligand (Cys) in the connecting loop and the fourth ligand (Cys) at the N-terminus of the second helix. The orientation of the helices is determined by the extensive hydrophobic packing and by the restraints imposed by the zinc-binding **(Figure 18B)** [110].

The structures of TAZ1 and TAZ2 are similar. There are only two main differences: TAZ1 has a 5 aa longer helix α1 and a different orientation of the helix α4 (which results in a shift in the position of one of the Zn^{2+} atoms) **(Figure 18C)** . The altered packing of helices in TAZ1 defines a different characteristic groove on the protein surface for ligand binding. It is likely that the structural differences between TAZ1 and TAZ2 contribute to their different substrate specificity and to their ability to target different sets of transcription factors [111].

Whereas the TAZ2 domain has been characterized as the interacting module for the best studied CH3 partners (p53 and E1A), the precise function of the ZZ domain remains unclear. The ZZ and the TAZ2 domain of CH3 fold independently and do not interact with each other. The structure of the ZZ domain (~ 45 aa) was solved by NMR and consists of two regions of β-sheets together with a small α-helix. This small domain contains two Zn^{2+} atoms coordinated with Cys_4 and Cys_2-His_2 motifs in a cross-brace fashion, which resembles the zinc-finger architecture **(Figure 19A-B)**. Given the similarity with RING

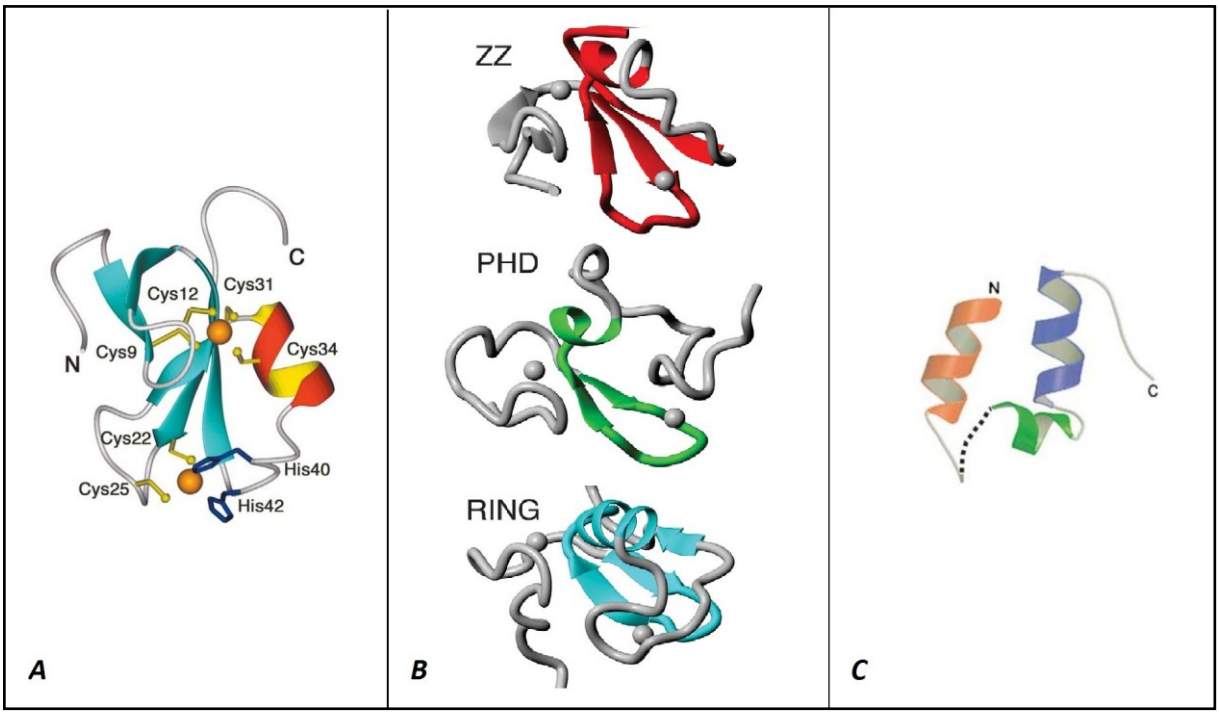


Figure 19 **Structures of ZZ and IBiD domains.** (A) NMR structure of the ZZ domain. Zinc ligands are shown [112]. (B) Structural comparison of interleaved zinc domains. The coloured backbone regions highlight the areas of similarity [112]. (C) NMR structure of the IBiD domain. Helix α_1 is shown in orange, α_2 in green and α_3 in cyan. The dashed line indicates four unassigned Gln residues [113].

domains and PHD fingers, it has been postulated that the ZZ domain may function as a ligand binding domain or as a protein scaffold [112].

The IBiD domain is a small structured region within the C-terminal QP-rich domain of p300/CBP which also mediates the binding of diverse proteins. In particular, it is involved in the interaction between IRF3 and p300/CBP in the context of the Type I interferon signalling. In analogy with the other functional domains of p300/CBP, the IBiD domain structure consists of three α-helices which pack together to form a hydrophobic core. The binding of ligand to this core induces fold both in the ligand and in the IBiD domain **(Figure 18C)** [113].

1.5.2. The p300/CBP 'core'

The central region of p300/CBP (residues 1045-1666 in p300) is highly conserved among multicellular organisms and comprises the catalytic domain (or HAT domain) and two adjacent effector modules (a BRD and a PHD finger). Sequence conservation and proteolysis studies suggest a conserved role of the whole central region (or 'core') for p300/CBP function (Data not published, from Daniel Panne in our lab) **(Figure 20)**.

As mentioned above, p300 and CBP share high sequence identity and they are commonly considered to be functionally equivalent. However, several studies suggest that p300 and CBP might also have some non interchangeable functions [114]. Some of these unique functions are due to small sequence divergences within the central region of the protein.

In the next sections, I will give an overview of what is structurally and functionally known about the central region of p300/CBP and I will mention some of the differences found between the CBP and p300 acetyltransferases. As the goal of my thesis was to characterize the central 'core' of p300 and the mechanisms of regulation of its AT activity, some of the following considerations are of pivotal importance for this work.

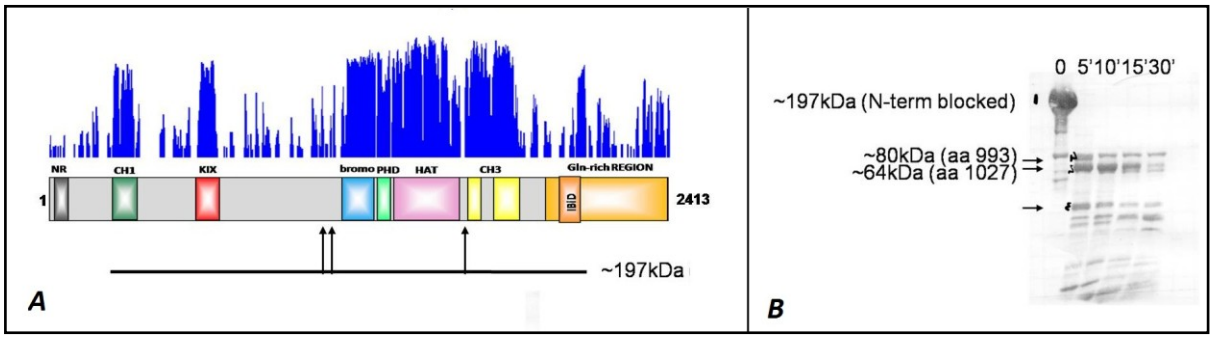


Figure 20 **Proteolysis studies on p300.** (A) Bar diagram of p300. Sequence conservation among higher eukaryotes is shown at the top. At the bottom, the construct tested (black line) and the 'core' fragments obtained (arrows). (B) SDS-page gel showing the fragments obtained during time-course limited proteolysis. Trypsin was used for the experiment. Arrows indicate resistant fragments. Incubation time is reported in minutes (') above the gel.Work by Daniel Panne, not published.

1.5.2.1. The structure of the p300 bromo- and PHD domains

The structure of the BRD of CBP (residues 1081-1196) in complex with different peptides containing mono-acetylated lysines was determined by NMR and X-ray crystallography **(Figure 21A)** [70, 115]. The structure adopts the conserved left-handed four-helical bundle typical of the BRD family (helices α_Z, α_A, α_B and α_C), with the ZA and BC loops forming the acetyl-lysine binding site. The conserved asparagine residue important for AcLys recognition is also present in CBP (N1168). The major difference with other BRDs lies in the presence of a two amino acid insertion (leucine 1120 and glycine 1121, L1120-G1121) within the ZA loop, which is involved in recognition of the AcLys and of the flanking residue (Kac+1) **(Figure 22)** [115].

This two amino acid insertion which confers specific substrate binding selectivity to CBP is also present in p300 **(Figure 22)**. However, it was demonstrated that substrate binding specificity can be diverse for the BRDs of the two proteins. For example, it was shown that both CBP and p300 can bind the transcription factor MyoD acetylated either on Lys99 or on both Lys99 and Lys102. However, the p300 BRD preferentially binds diacetylated MyoD peptides, whereas CBP BRD distinguish poorly between mono- and di-acetylated ones [116]. The fact that p300 can bind diacetylated substrates was also suggested by sequence comparison with the Brdt-BD1 bromodomain discussed above [68].

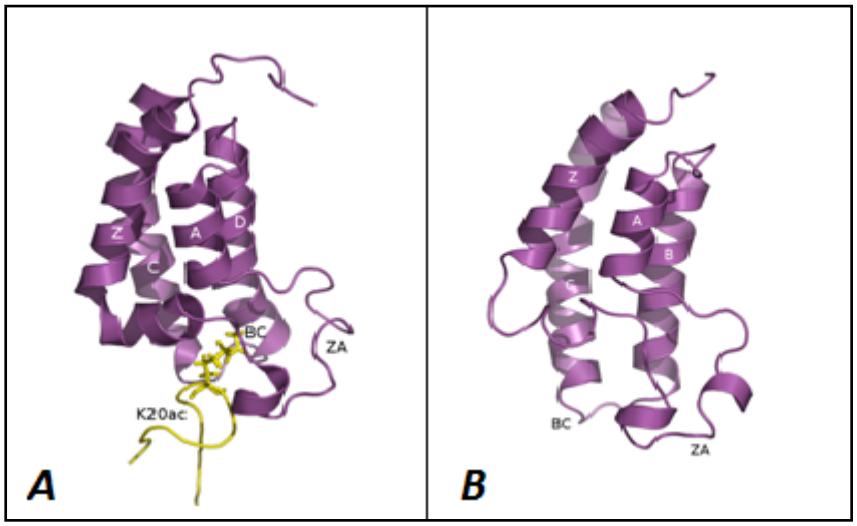


Figure 21 **Bromodomains of CBP and p300.** (A) NMR structure of the CBP bromodomain in complex with H4K20ac peptide (yellow sticks). (B) Crystal structure of the ligand-free bromodomain of p300. Loop ZA and BC are shown.

The structure of a peptide-free p300 bromodomain (residues 1047-1158) was solved by X-ray crystallography (PDB3I3J) and showed the typical bromodomain fold **(Figure 21B)**. Nothing is known about the mode of substrate recognition of the p300 bromodomain, apart from the presence of the conserved asparagine, N1131 **(Figure 22)**.

The structure of the PHD domain of p300/CBP is not available to date. It is worth mentioning that the PHD domain of p300/CBP is atypical and might have different properties when compared to other PHD fingers. Indeed, it contains a large insertion within the loop1, which comprises 42 residues instead of the usual 8-13 amino acids **(Figure 23)** [10]. Moreover, structural divergence between p300 and CBP is suggested by some functional studies carried out on the two proteins. Indeed, the PHD finger is essential for AT

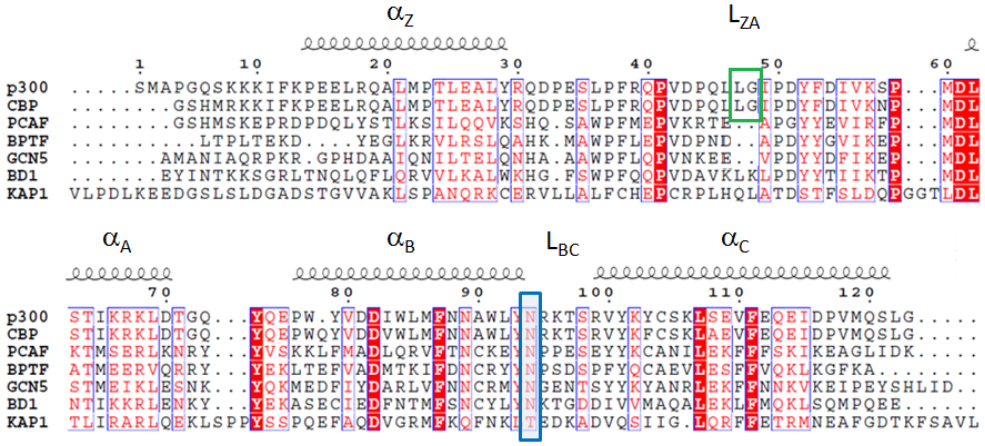


Figure 22 **Sequence alignement of bromodomains.** The conserved asparagine is shown in the blue box. The two aminoacid insertion (L1120-G1121) is highlited by the green box.

activity of CBP and is an integral part of the HAT domain [117]. In contrast, the PHD finger of p300 has been shown to be dispensable for enzymatic activity and is not part of the HAT domain [118].

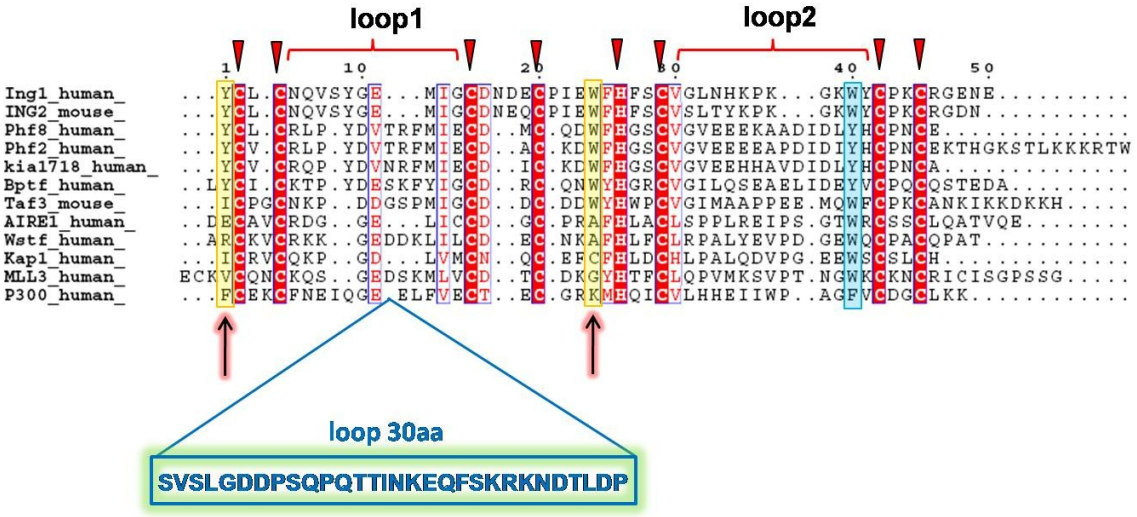


Figure 23 **Sequence alignement of PHDs.** The conserved residues coordinating the zinc-atoms are indicated by red triangles. The residues important for methyl-lysine recognition are indicated by pink arrows. The 30 aminoacid insertion is reported in the blue box.

1.5.2.2. The bromo-PHD module

Many studies have been carried out to decipher the role of the p300 BRD. In experiments performed with HDAC inhibitors and cells containing bromo-deficient p300, it was shown that the p300 BRD is important to maintain the basal level of histone acetylation and to induce transcriptional activation of p300-dependent genes [119]. Other studies demonstrated that the p300 BRD serves as substrate targeting module because it was found to be involved in the stable interaction of p300 with acetylated transcription factors and chromatin [115, 116, 120]. However, the p300 BRD might not be the only player in substrate recognition. For example, it was demonstrated that the p300 BRD is necessary but not sufficient for the binding of p300 with chromatin [120].

An hypothesis is that the PHD finger of p300 could be as well involved in binding to chromatin, forming a functional module with the adjacent BRD (a bromo-PHD module). In nucleosome retention assays and EMSAs, Ragvin *et al.* [88] reported that the p300 PHD finger cooperates with the BRD to confer a robust association with hyperacetylated chromatin. Moreover, the replacement of the p300 PHD finger by other canonical PHD fingers resulted in loss of nucleosome binding activity [10]. As binding of the p300 PHD

finger to chromatin is acetylation independent, Ragvin *et al.* also proposed that the PHD domain may be involved in the binding to another part of the targeted nucleosome or to both nucleosomes and other proteins simultaneously [88].

In conclusion, the BRD and the PHD finger in p300 may act as other well-known linked binding modules and hence be involved as a unit in substrate recognition or display novel properties, such as regulatory functions.

1.5.2.3. The HAT domain

The structure of the p300 HAT domain (residues 1284-1666) in complex with a potent inhibitor (LysCoA) was solved by X-ray crystallography at 1.7Å resolution **(Figure 24A)** [92]. The structure is made up of a β-sheet composed of seven β-strands surrounded by nine α-helices and several loops. The last β-strand and the last three α-helices form the so-called C-subdomain. The central region which binds to the acetyl-CoA cofactor (corresponding to the A, B and C motifs of the GNAT superfamily) shows structural conservation with other HATs families **(Figure 13A)**. In contrast, the structural elements flanking this central region are divergent from other HATs [92].

The presence of a long loop (L1) is unique to p300. This loop is intimately associated with the CoA portion of the LysCoA inhibitor and contributes to one-third of the protein-cofactor interactions. As the loop L1 covers one side of the cofactor, the AcCoA binding site is more buried in p300 than in other HATs. In addition, comparison of the electrostatic surface potential of the substrate binding surfaces of the HAT domains of Gcn5, Esa1 and p300 shows divergence between p300 and other HATs. Gcn5 and Esa1 are characterized by deep and apolar binding pockets, whereas p300 has a shallow and more electronegative substrate binding site **(Figure 13 B-D)** [92]. This is consistent with the different binding properties of p300 HAT, which prefers substrates with multiple positive charges, compared to other HATs [106]. Unlike other HATs, which have limited specificity, p300 HAT is able to acetylate the four histones (*in vivo* and *in vitro*) and many non-histone proteins, such as transcription factors [3, 106].

There are two shallow and electronegative pockets corresponding to the substrate binding site of p300 (pockets P1 and P2). A narrow, shallow and electronegative groove connects P1 and P2 **(Figure 24B)**. P1 is the site which accommodates the lysine substrate and the AcCoA cofactor, whereas the pocket P2 is located 10Å away from P1 and possibly contributes to the binding of protein substrates. Indeed, p300 substrates contain

preferentially Lys or Arg chains at position -/+3 or -/+4 of the primary acetylation site (about 10Å away) [92, 106].

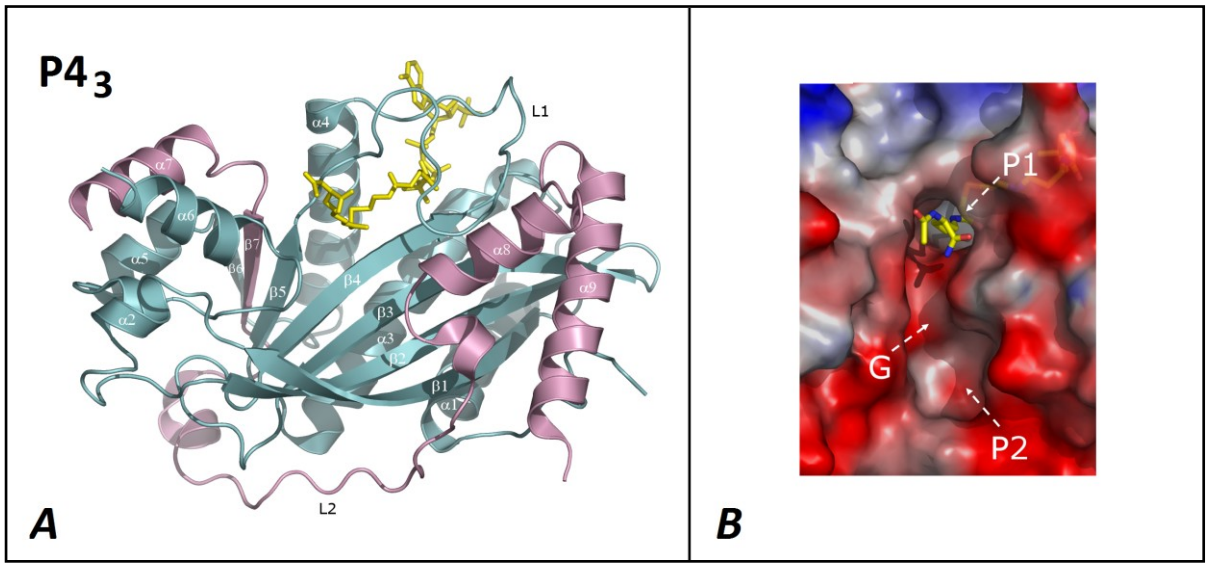


Figure 24 **Structure of the p300 HAT domain.** (A) Crystal structure of the p300 HAT domain. N and C subdomain are coloured in cyan and pink respectively. The LysCoA inhibitor is shown in yellow sticks. (B) Electrostatic surface representation of the putative substrate binding site of p300 HAT. Pocket 1, pocket 2 and the connecting groove are annotated as P1, P2 and G, respectively.

1.5.2.4. HAT activity and regulation

The structure of the HAT domain and biochemical studies has elucidated the mechanism of catalysis of p300/CBP. As discussed above, p300/CBP does not need the formation of a stable ternary complex for catalysis and promotes the acetyl-transfer in a way consistent with a Theorell-Chance (or 'hit and run') mechanism **(Figure 12)**. Moreover, p300/CBP does not critically depend on the presence of a specific Asp/Glu residue as a base for catalysis, unlike the GNAT and MYST superfamilies [92, 121].

Structure-based mutagenesis studies have identified two other critical residues for acetyl transfer catalysis in p300, the tyrosine (Tyr) 1467 and the tryptophan (Trp) 1436. Tyr1467 is likely to play an important orienting role as well as a general acid function in protonating the CoA leaving group. On the other hand, Trp 1436 plays a role in guiding the substrate lysine side chain into the binding site. Substitutions of Tyr1467 with a Phe and of Trp1436 with an Ala greatly reduced HAT activity [92].

In addition to the biochemical and mutagenesis studies, a lot of effort has been done in order to characterize p300 HAT regulation. Studies in this direction are very important since misregulation of p300/CBP has often been associated with the development of severe diseases [5]. An interesting discovery about p300 HAT regulation was done few years ago by Thompson and colleagues [6]. They reported that the p300 HAT domain contains a highly basic loop which is responsible of p300 autoregulation **(Figure 25)**. This loop can function as an autoinhibitory loop when hypoacetylated, locking the enzyme in an autoinhibited conformation and possibly impairing the access of substrates to the active site. Autoacetylation within the loop switches the enzyme to its active form and stimulates HAT activity. This mechanism of autoregulation is conceptually analogous to the mechanism of activation of kinases by autophosphorylation **(Figure 16)** [6].

The lysine-rich autoinhibitory loop was proteolytically cleaved in the structure of the HAT domain by Liu *et al.* [92]. However, it was proposed that the basic autoinhibitory loop, when hypoacetylated, could be accommodated in the highly electronegative substrate binding pocket (P2) of p300 and impede the access of substrates to the p300 catalytic site **(Figure 24B)** [92]. Moreover, kinetic and mass spectrometric analysis of p300 HAT autoacetylation demonstrated that autoacetylation within the autoinhibitory loop was preferentially intermolecular (trans-autoacetylation) [7]. Kinetic analysis also revealed that rapid autoacetylation of approximately five-seven sites, several of functional importance, occurs first and is then followed by a slower and more global acetylation. Overall, up to 17 residues were found to be autoacetylated within the p300 autoinhibitory loop [7, 92].

Figure 25 **Sequence alignment of the p300/CBP autoinhibitory loop region.** Some of the lysines that are autoacetylated within the loop are highlighted in orange.

1.5.3. p300/CBP: implication in diseases

Germline and somatic mutations of p300 and CBP, as well as chromosomal translocations, were associated with p300/CBP misregulation and the development of severe diseases [5, 122].

The structure of the p300 HAT domain allowed rationalization of several p300/CBP inactivating mutations that have been observed in various cancers [92]. Mutation of the catalytic residue Trp1436 has been associated with the Rubinstein-Taybi Syndrome, whereas point mutations destabilizing the loop L1 were correlated to other severe forms of cancer (i.e. D1399Y and G1375E substitutions in primary colon cancer and in lung cancer, respectively). From structural interpretation, all these mutations affect critical catalytic residues in the HAT domain and severely affect HAT activity [92]. The structure of the p300 HAT domain is certainly important for understanding the molecular basis of disease development but it is also instrumental for structure-based drug design of more specific therapeutic compounds.

It is worth mentioning that the majority of point mutations, translocations and missense or in-frame deletions of p300/CBP associated to the development of cancer are mapped in the 'core' region of the protein **(Figure 26)** [5, 122]. For instance, PHD finger mutations can lead to loss of CBP acetyltransferase activity and are alone sufficient for the development of the Rubinstein-Taybi Syndrome [2]. This suggests that the BRD and PHD finger of p300/CBP, and not only the HAT domain, are important for protein function. The structural and biochemical characterization of the whole central 'core' of p300/CBP is pivotal for understanding p300/CBP regulation and for more accurate structure-based drug design.

Several lines of evidence also support the model that p300/CBP acts as a tumour suppressor [5]. For example, p300/CBP can modulate at multiple levels the p53-mediated response to DNA damage. Loss of p300/CBP leads to disruption of p53 activation, reduced p53 stability and decreased ability to promote transactivation of target genes. Upon DNA damage, inhibition of p53-mediated apoptosis due to p300/CBP loss can have severe consequences on cell fate [5].

Sequestration of p300/CBP and consequent p53 inactivation can result in oncogenesis. A well-studied example is the oncogenic fusion protein Brd4-NUT, which is involved in the appearance of a highly aggressive tumour called NUT midline carcinoma (NMC) **(Appendix)** [123]. Brd4-NUT results from a chromosomal translocation which fuses the NUT gene (Nuclear protein in testis) on chromosome 15 to the Brd4 gene (a

double BRD containing protein) on chromosome 19. The NUT moiety of the fusion protein can interact with and activate p300. Brd4-NUT sequentially recruits p300 though its NUT moiety in order to form hyperacetylated and transcriptionally inactive chromatin foci. Sequestration of p300 into these foci leads to p53 inactivation, which results in malignant cell transformation **(Appendix)** [123].

Sequestration of p300/CBP is a general mechanism for development of severe diseases other than cancer. In neurodegenerative diseases, such as the Huntington's disease, p300/CBP function is inhibited because of the presence of expanded polyglutamine-repeats in key proteins (i.e. the hungtingtin). These repeats can bind avidly to the HAT domain and the C-terminus of p300/CBP and inhibit enzyme activity. Sequestration of p300/CBP by polyglutamine-repeats containing proteins is the major cause of neurotoxicity in neurodegenerative diseases [122].

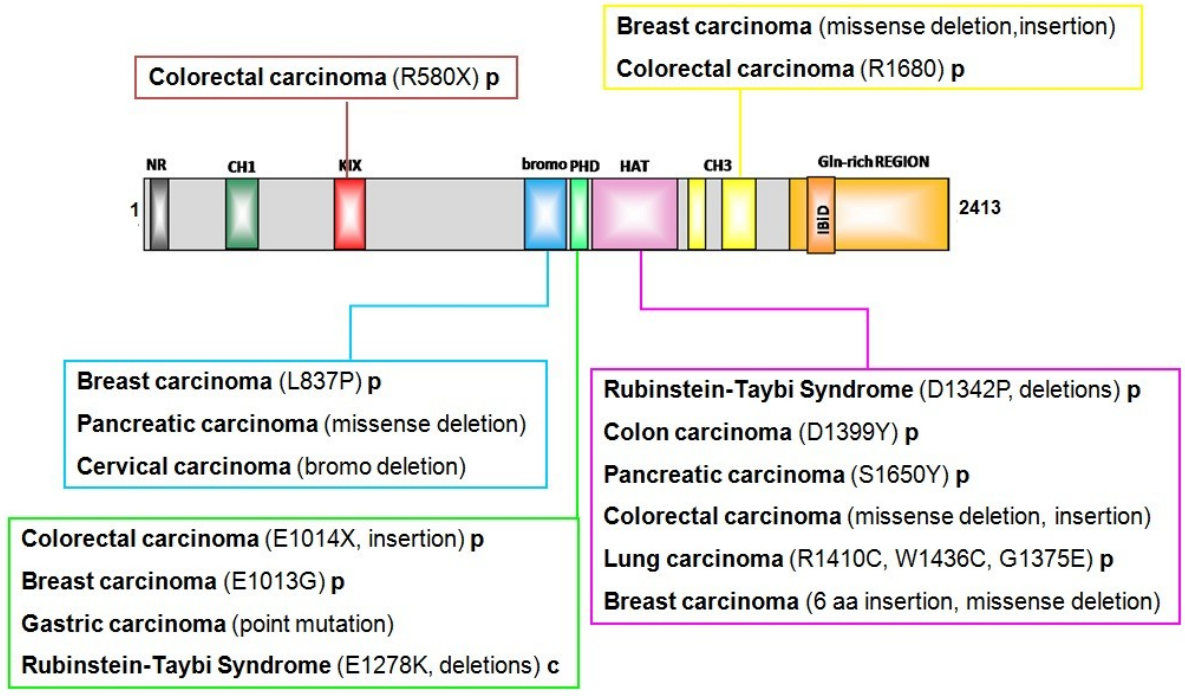


Figure 26 **p300/CBP and cancer.** Point mutations, deletions and insertions linked to diverse forms of cancers are indicated in the boxes. Residue numbers for the point mutations refer to p300 (**p**) or CBP (**c**).

1.6. THE SCIENTIFIC QUESTIONS

The p300 acetyltransferase is a very important transcriptional coactivator, whose misregulation has severe impact on cellular processes. This renders the study of p300 regulation very important for medical research. Autoregulation is thought to occur within the central catalytic 'core' of the protein. This 'core' contains the histone acetyltransferase domain (or HAT) and the module bromo-PHD (BP). Autoacetylation of a lysine rich loop (autoinhibitory loop) within the HAT domain was reported to regulate protein activity [6].

During my thesis work, I tested the hypothesis if p300 activity is regulated by the autoinhibitory loop and the bromo-PHD (BP) module in a similar manner to that of activation loop and SH2-SH3 domains in kinases. The idea is that acetylation/deacetylation events in p300/CBP may play a role similar to phosphorylation/dephosphorylation in regulation of protein kinase activity [8]. In particular, I focused my studies on some major scientific questions.

- Is the the BP module involved in HAT regulation? Which is the role of the BP module in this regulation?
- Is the BP module involved in substrate targeting? What is the substrate specificity of the BP module?
- Can the structures of the inhibited and active conformations of the p300 'core' explain the mechanism of p300 regulation and inhibition?

The first aspect I have investigated was if the BP module could stabilize the inactive autoinhibited conformation of the p300 'core'. This autoinhibited conformation is characterized by a hypoacetylated state, in particular within the autoinhibitory loop [6]. The BP module may play a role which is acetylation dependent, since it contains an acetyl-lysine binding domain (the BRD). In order to test my hypothesis, I performed *in vitro* biochemical analysis with enzymatic assays, limited proteolysis, ITC binding studies, analytical size exclusion chromatography and pull-downs.

Moreover, the BP module was already shown to be able to target p300 to specific substrates [88, 115]. However, its binding specificity was never described. Therefore, I focused part of my work in this direction. I carried out ITC experiments using modified histone peptides. The idea was that the BP module could bind substrates which are already modified by p300. This would allow the sequential recruitment of p300/CBP molecules to

the same region of chromatin and lead to the formation of positive feedback loops of acetylation.

Finally, the goal of my thesis was to structurally characterize the central 'core' of the p300 acetyltransferase. I studied the p300 'core' using X-ray crystallography and small angle X-ray scattering (SAXS). The advantage of having a structure of p300 is that it allows a better understanding of the mechanism of regulation of its enzymatic activity. Moreover, analyzing the structure of p300 can be very informative for medical research. P300 can be an important target for therapies against several forms of disorders, such as the Rubinstein-Taybi Syndrome, for which specific and effective treatments are not available to date.

2. MATERIALS AND METHODS

RÉSUMÉ EN FRANÇAIS

Dans le chapitre matériel et méthodes sont d'abord décrites les techniques de clonage, d'expression et de purification des protéines caractérisées dans le cadre de ce projet de thèse. Plus particulièrement, l'expression des protéines hétérologues a été réalisée dans *E. coli* BL21 (DE3) ou dans des cellules d'insectes Hi5. Toutes les protéines ont été purifiées à l'homogénéité et leur pureté a été contrôlée par électrophorèse sur gel de polyacrylamide en présence de SDS et par spectrométrie de masse (MS).

Les méthodes biochimiques, biophysiques et structurales qui ont été utilisées pour caractériser les macromolécules sont décrites dans ce chapitre. Il s'agit notamment de l'analyse des protéines d'intérêt après digestion partielle par la trypsine, de tests qualitatifs de l'activité (détection des niveaux d'acétylation par western blot) et quantitatifs (détection de l'activité par un test colorimétrique, en utilisant DTNB). Il s'agit aussi des tests d'interaction entre le module BP et les queues des histones ou le domaine HAT (GST-*pull-downs*, tests en chromatographie d'exclusion, calorimétrie isotherme à titration ou ITC) ainsi que d'une caractérisation structurale du 'noyau' de p300/CBP (cristallographie par diffraction des rayons X et diffusion des rayons X aux petits angles ou SAXS).

2. MATERIALS AND METHODS

2.1. CONSTRUCTS

A list of the constructs characterized during this thesis work is reported in **Table 1**. Bromo-PHD constructs were cloned in pETM-33 or pETM-30 vectors (EMBL). These vectors contain a GST sequence (221 aa) and a TEV recognition site (ENLYFQG) upstream of the MCS (multiple cloning site), allowing for cloning of N-terminal GST-tag fusion proteins with cleavable tag. The p300 HAT domain constructs were cloned in the pET-21a vector (Novagen), for the expression of N-terminal His-tag (polyhistidine, HHHHHH) fusion proteins. The His-tag can be cleaved off using TEV protease. The p300 'core' constructs were cloned both in pETM33 and pFASTBAC1 (Invitrogen) vectors as GST-fusion proteins, TEV cleavable. Attempts to use different tags, such as FLAG tag (DYKDDDK) or STREP tag (WSHPQFEK), to express longer/shorter constructs and to clone in other vectors such as pETDuet-1 (Novagen) were also made but are not discussed here. Point mutations were obtained using the Quick Change® Site-Directed Mutagenesis Kit from Stratagene. The His-SIRT2 construct in pFASTBAC1 vector was a gift from the Carey lab (University of California, LA).

Table 1. **List of constructs used during the thesis work.**

Construct	Residues	Vector	Tag	Mutations and variants
p300 'core'	1048-1666	pETM-33, pFASTBAC1	GST	*wt*, Y1467F
p300 'core' Δloop	1048-1666	pETM-33, pFASTBAC1	GST	Δ1520-1580 (linker SGGSG), /Y1467F
p300 bromo-PHD	1048-1276	pETM-33	GST	*wt*, N1131A, K1253A
CBP bromo-PHD	1085-1313	pETM-30	GST	*wt*
p300 HAT	1284-1666	pET-21a	His	*wt*, Y1467F
p300 HAT Δloop	1284-1666	pET-21a	His	Δ1520-1580 (linker SGGSG), /Y1467F
SIRT2	Full-length	pFASTBAC1	His	*wt*

2.2. EXPRESSION AND PURIFICATION

2.2.1. Expression and purification of GST-tagged proteins from *E. coli*

The GST fusion proteins were expressed in *E. coli* BL21 (DE3) using LB medium enriched with 100 μM $ZnCl_2$. Cells were grown under agitation at 37°C until they reached an OD_{600} of ~0.6. Protein expression was induced with IPTG (1 mM) at 17°C overnight (o/n). Cell pellets were resuspended in lysis buffer (20 mM Tris/HCl pH 8.0, 300 mM NaCl, 1 mM DTT, 5 μM $ZnCl_2$ and one tablet of Complete Protease Inhibitors EDTA-Free from Roche Applied Science) and lysed using a Microfluidizer (from Microfluidics). The lysate was cleared by centrifugation at 20000 rpm for 40 minutes and the supernatant was applied to a Glutathione Sepharose 4 Fast Flow resin (GE Healthcare). The resin was washed with wash buffer (20 mM Tris/HCl pH8.0, 300 mM NaCl, 1 mM DTT and 5 μM $ZnCl_2$) and incubated with TEV protease (1:100 w/w) o/n at 4°C. The protein sample was then washed off the column using wash buffer and loaded onto an IMAC Sepharose 6 Fast Flow resin (GE Healthcare) pre-charged with Ni^{2+} ion and pre-equilibrated in wash buffer. The His-TEV protease was removed from the protein sample during the His affinity purification step. The untagged protein was further purified by gel filtration on a High Load 16/60 Superdex 75 (or 200) column (GE Healthcare) equilibrated in 20 mM Hepes, 300 mM NaCl, 0.5 mM TCEP and 5 μM $ZnCl_2$. The final protein was concentrated to 10 mg/ml and SDS-PAGE gels were run to verify purity of the sample. Aliquots were flash frozen in liquid N_2 and stored at -80°C.

2.2.2. Expression and purification of His-tagged proteins from *E. coli*

The His-tagged constructs were expressed in *E. coli* BL21 (DE3 or DE3/Rosetta) using autoinducing media ZYP-5052 [124]. Cells were grown at 37°C until they reached an OD_{600} ~0.6 and then temperature was decreased to 17°C o/n. Cell pellets were resuspended in lysis buffer (50 mM Tris/HCl pH 8.0, 500 mM NaCl, 20 mM Imidazole, 1 mM DTT and one tablet of Complete Protease Inhibitors EDTA-Free from Roche Applied Science) and lysed using a Microfluidizer (Microfluidics). The lysate was cleared by centrifugation at 20000 rpm for 40 minutes and the supernatant was applied on a 5-ml His-TRAP™ FF crude column (GE Healthcare) pre-charged with Ni^{2+} ion. The column was washed with wash

buffer (50 mM Tris/HCl pH 8.0, 500 mM NaCl, 20 mM Imidazole and 1 mM DTT) and protein was eluted in a single step using elution buffer (wash buffer containing 500 mM Imidazole). The His-tag was cleaved off with TEV protease (1:100 w/w) in dialysis o/n at 4°C using 2 L of wash buffer while stirring. The His-tag and the His-TEV were then removed from the preparation with a second His-affinity purification step. Untagged protein was then further purified by gel filtration on a High Load 16/60 Superdex 200 column (GE Healthcare) equilibrated in 20 mM Hepes, 300 mM NaCl and 0.5 mM TCEP. The final protein was concentrated to 10 mg/ml and SDS-PAGE gels were run to verify purity of the sample. Aliquots were flash frozen in N_2 and stored at -80°C.

2.2.3. Baculovirus cultivation and protein purification from insect cells

For constructs cloned in pFASTBAC1 vectors, recombinant baculoviruses were propagated in Sf21 cells with Sf900-II SFM medium (Invitrogen). For protein production, Hi5 cells in Express Five SFM medium (Invitrogen) were infected with recombinant baculoviruses at a multiplicity of infection of 2.0, maintained in shake flasks at 28°C and harvested by centrifugation 72h postinfection. Cell pellets were lysed by freezing and then thawed in lysis buffer (20 mM Tris/HCl pH 8.0, 300 mM NaCl, 0.5 mM TCEP, 5 μM $ZnCl_2$, one tablet of Complete Protease Inhibitors EDTA-Free from Roche Applied Science and 10 U of Benzoase). The cells were homogenised by sonication and the lysate was clarified by centrifugation at 20000 rpm for 1h. The clarified lysate was applied to a Glutathione Sepharose 4 Fast Flow resin (from GE Healthcare) and purification proceeded as described for the GST-tagged proteins from *E. coli* (excluding the His-purification step).

2.3. TRYPTIC DIGESTS

Trypsin from Promega (V5280) was used for time-course tryptic digest experiments. Proteins were diluted to 1 mg/ml in buffer containing 20 mM Tris/HCl pH 8.0, 300 mM NaCl and 1 mM DTT. Digest was started by the addition of diluted trypsin to the protein mix (molar ratio 1:400 of trypsin:protein was used in most of the experiments). Samples were taken at several time points. At each time point, trypsin was inactivated by addition of SDS-loading dye and boiling of the sample for 10 min. Protein fragments were then visualized by SDS-PAGE Gel Electrophoresis and Coomassie Blue staining. In this experiment, 12% polyacrylamide gels were used.

2.4. WESTERN BLOT

2.4.1. HAT assay

HAT assay were performed using full-length histones H3 and H4 from *C. elegans* as p300 substrates. The p300 and the histone substrates were diluted in reaction buffer (20 mM Hepes pH 7.9, 300 mM NaCl, 50 μg/ml BSA, 5 mM DTT and 100 μM AcCoA) at a molar ratio of 1:10 (enzyme:substrate). The reaction mix was then incubated at 30°C for 30 minutes. The reaction was stopped by addition of SDS-loading buffer and boiling of the samples for 10 minutes. Aliquots were loaded on 12% or 15% polyacrylamide SDS-PAGE gels.

Lysine acetylation was detected by Western Blot (buffers used are indicated in **Table 2**). After SDS-PAGE, the protein bands were transferred on nitrocellulose membrane using a semi-dry transfer cell apparatus (Bio-Rad Trans Blot SD) according to instructions from the manufacturer. After transfer, the membrane was blocked for 1h, washed with wash buffer and then incubated o/n in a solution containing anti-acetyllysine antibody (primary antibody generated in rabbit, 1:1000, Cell Signaling Technology, 9441S) at 4°C. After incubation with the primary antibody, the membrane was washed and incubated with an anti-rabbit IgG HRP-linked antibody (secondary antibody, 1:2000, Cell Signaling Technology, 7074S) for 1 h at room temperature (RT). For detection, the membrane was incubated with LumiGlo solution (Cell Signaling Technology, 7003S) containing peroxide (Cell Signaling Technology, 7003S) and exposed to an X-ray film for 5-10 seconds. The film was developed using an AGFA CURIG 60 X-ray film processor.

2.4.2. Deacetylation assay

Deacetylation assays were performed using p300 'core' constructs as SIRT2 substrates. The reaction buffer used for the deacetylation reaction contained 20 mM Hepes pH 7.5, 300 mM NaCl, 0.5 mM TCEP, 5 mM $MgCl_2$ and 1 mM NAD+. Proteins were mixed in a 1:10 (SIRT2:p300) molar ratio and diluted in reaction buffer to a final concentration of 1 mg/ml for p300. The mix was incubated at room temperature (RT) and samples were taken from the reaction mix at different time points (from 30 min to o/n). At each time point the reaction was stopped by addition of SDS-loading buffer and the samples

were boiled for 10 minutes. Acetylation was monitored using the same assay as described above.

Table 2. Buffer used for western blot.

Buffer	Composition
TBS	20 mM Tris/HCl pH 7.6, 140 mM NaCl
Blocking buffer	TBS, 0.1% Tween-20, 5% w/v non fat dry milk
Wash buffer	TBS, 0.1% Tween-20
Antibody dilution buffer	TBS, 0.1% Tween-20, 5% BSA

2.5. GST PULL-DOWNS

GST pull-downs were carried out with recombinant GST-tagged bromo-PHD (BP) constructs from *E.coli* BL21 (DE3). GST-BP was purified as described in chapter 2.2.1. The only difference was that the TEV cleavage step was replaced by an elution step of the GST-tagged protein with wash buffer containing 10 mM reduced glutathione.

For the pull-downs, GST-BP and the putative binding partners were pre-incubated in binding buffer (20 mM Tris/HCl pH8.0, 300 mM NaCl, 0.5 mM TCEP and 5 μM $ZnCl_2$) at equimolar ratios. Glutathione Sepharose 4 Fast Flow resin (from GE Healthcare) was pre-equilibrated in the same binding buffer and resuspended in a 50% w/v solution. The resin was then added to the protein mix (10 μl final resin for 100 μg final GST-BP). Incubation with the resin was performed o/n on a rotating wheel at 4°C. After incubation, the resin was pelleted and the flow-through (FT) was removed. The resin was then washed three times with wash buffer (binding buffer containing 0.1% Tween-20). Elution was done by addition of 1xSDS loading buffer to the resin and the samples were boiled for 10 minutes. Protein samples from the pull-downs were run on a 12% polyacrylamide gel and visualized by Coomassie Blue staining.

2.6. ANALYTICAL SIZE EXCLUSION CHROMATOGRAPHY

Size exclusion chromatography analytical runs were performed using GST-cleaved BP and His-cleaved HAT domain. Samples (100 μl each) of the proteins alone and of an equimolar mix of the two were equilibrated in S200 buffer (20 mM Hepes pH 7.5, 300 mM

NaCl, 0.5 mM TCEP and 5 μM $ZnCl_2$) and incubated on ice for one hour. After incubation, the samples were injected on a Superdex 200 10/300 GL column (GE Healthcare) equilibrated in the same S200 buffer. Column calibration was performed using a gel filtration standard (Bio-Rad, 151-1901).

2.7. ISOTHERMAL TITRATION CALORIMETRY (ITC)

Isothermal titration calorimetry (ITC) was performed using a high-precision VP-ITC or iTC200 systems (MicroCal Inc., Nothampton, MA). Proteins were extensively dialyzed against titration buffer (20 mM Hepes pH 7.5, 300 mM NaCl, 0.5 mM TCEP). Peptide solutions were prepared using the same dialysis buffer. All the solutions were filtered and degassed to avoid bubble formation and they were equilibrated to the working temperature (25°C) before each experiment. Protein solutions at about 50-70 μM were titrated with the different peptide or protein solutions at appropriate concentrations (ranging from 0.5 mM to 5 mM) in order to define the titration curves. Titrations were carried out by injection of a constant volume of peptide (7 μl for the VP-ITC and 1.5-2 μl for the iTC200). Experiments were performed at least in duplicates. The heat released/absorbed after each injection could be obtained from the integral of the calorimetric signal. The resulting binding isotherms were analyzed by non-linear least-squares fitting of the experimental data to a model corresponding to a single set of identical sites. The corresponding equation is the following:

$$Q = \frac{n[M_t]\Delta H V_0}{2}\left[1 + \frac{[L_t]}{n[M_t]} + \frac{1}{nK_a[M_t]} - \sqrt{\left(1 + \frac{L_t}{n[M_t]} + \frac{1}{nK_a[M_t]}\right)^2 - \frac{4[L_t]}{n[M_t]}}\right]$$

where Q is the net heat of binding, n is the number of binding sites, K_a is the association constant, V_0 is the active cell volume, ΔH is the enthalpy of binding, and $[M_t]$ and $[L_t]$ are the total concentration of the protein in the cell (macromolecule) and the ligand, respectively. In the analysis, the number of binding sites, the association constant and the binding enthalpy were considered as variables. Analysis of the data was performed using the MicroCal Origin® program (OriginLab Corporation, Northampton, MA).

2.7.1. What is ITC and how does it work?

ITC is a physical technique used to measure biomolecular interactions. When substances bind, heat is either generated or absorbed. ITC directly measures the heat changes during a binding event. This initial measurement allows for direct determination of all binding parameters (n, K_a, ΔH and ΔS or entropy change) in a single experiment, thereby providing a complete thermodynamic profile of the binding process. The parameter determination is based on basic thermodynamic relationships:

$$\Delta G = - RT \ln K_a = \Delta H - T\Delta S$$

Where ΔG is the Gibbs energy change, R is the gas constant and T is the absolute temperature. For a more detailed description, see Velazquez Campoy *et al.* (2005) [125].

An isothermal titration calorimeter is composed of two identical cells (reference and sample cell) mounted in a cylindrical adiabatic environment and connected to the outside through narrow access tubes **(Figure 27)**. Sensitive thermoelectric devices are used to detect temperature differences between the reference cell (filled with buffer or water) and the sample cell containing the macromolecule. During the experiment, a feedback control system supplies thermal power continuously to maintain the same temperature in both cells.

Briefly, the macromolecule solution is located inside the sample cell and the ligand solution in the injector syringe. When ligand is injected into the sample cell, heat is

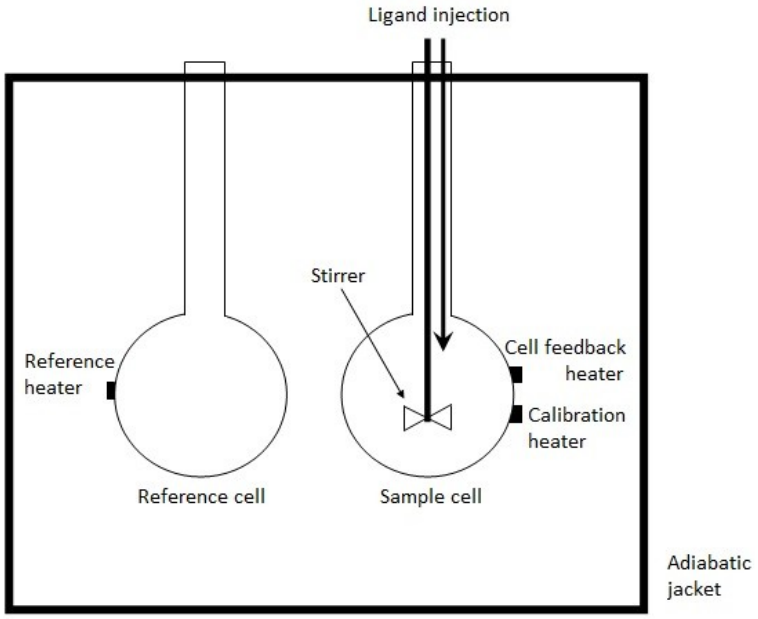


Figure 27 **Schematic representation of an ITC instrument.** The syringe rotates in place during the ITC experiment to provide continuous mixing in the sample cell. For details about the instrument setup see in the text.

generated or absorbed in direct proportion to the amount of binding. The ligand solution is titrated periodically into the sample cell and the temperature feedback control system modulates the power supplied to maintain temperature imbalance between the sample and the reference cells. The heat associated with each ligand injection is proportional to the increase in concentration of the complex and is calculated by integrating the area under the deflection of the signal measured (amount of heat per unit of time provided to maintain both cells at the same temperature). As the macromolecule in the cell becomes saturated with ligand, the heat signal diminishes until only the background heat of dilution is observed. Analysis of the fitting of the experimental data to an appropriate model using non-linear regression allows for the determination of all the binding parameters [125].

2.8. SMALL ANGLE X-RAY SCATTERING (SAXS)

SAXS experiments were carried out at the SWING beamline (SOLEIL Synchrotron, Paris) or on ID14-3 (ESRF). At SWING, the SAXS sample chamber was connected online to an HPLC (Agilent 1200 series) and a gel filtration column, in order to separate different molecular species prior to SAXS data collection. This was useful to discriminate between protein aggregates, oligomers and monomers in the data analysis. The beamline wavelength was set up at 0.99 Å and the scattering was collected on an AVIEX170170 CCD detector under vacuum. This setup allowed the measurement of the scattering vector ($q=4\pi\sin(\theta)/\lambda$, where 2θ equals the scattering angle) within the q-range $0.0005<q<0.62$.

All samples were concentrated to 15 mg/ml and centrifuged at 13000 rpm for 15 minutes before injection. The HAT domain sample was pre-incubated with 3-fold LysCoA (a p300 HAT inhibitor) for 30 minutes on ice before centrifugation. 50 µl of each sample were injected onto the gel filtration column equilibrated in buffer (20 mM Hepes pH7.5, 300 mM NaCl, 0.5 mM TCEP, 5 µM $ZnCl_2$) and directly eluted (200 µl/min) into a quartz capillary of an inner diameter of 1.5 mm (SAXS sample chamber). Data collection was done throughout the elution time (from 11 to 20 minutes) and 200-250 frames were collected for each sample elution using an exposure time of 2 seconds. Only selected frames corresponding to the main elution peak were taken into account for the analysis. Buffer measurements were preformed immediately prior the protein elution peak (100 frames).

Each frame was normalized to the transmitted intensity. Radial averaging was carried out using the program *ActionJava* developed at the SWING beamline. Scaling of the data was achieved using water as a reference. Buffer subtraction was performed as a routine

in MATLAB®, which enables the handling of a large amount of data. Average protein concentration was estimated from the UV absorption at 280 nm using a spectrometer located upstream of the SAXS sample cell. All subsequent data analysis and modelling were carried out with PRIMUS and other programs of the ATSAS suite (for reviews see [126, 127]).

From the corrected scattering curves, the forward scattering I(0) and the radii of gyration (R_g) were obtained using the Guinier approximation assuming that, at very small angles ($q<1/3R_g$), intensities can be represented as $I(q)=I(0)\exp(-(qR_g)^2/3)$. These parameters were also computed from the entire scattering patterns using the program GNOM, which provides the maximum dimension (D_{max}) and the distance distribution function (p(r)) of the particle. The molecular masses of the solutes were estimated from the forward scattering by normalization against water (the reference solution). *Ab-initio* models were obtained using DAMMIN and DAMAVER. Rigid-body refinement against the scattering data was done using SASREF. Comparison with atomic models and high-resolution structure validation was done using CRYSOL.

2.8.1. What is SAXS and how does it work?

SAXS is a technique that is used for the structural characterization of solid and fluid materials in the nanometre range. Structural information can be obtained in a resolution range between 5 and 25 nm, and of repeat distances in partially ordered systems of up to 150 nm in size. In a SAXS experiment, macromolecules in solution are exposed to X-ray and the scattering is recorded at very low angles (0.1-10°) on a detector. The scattering intensity I(q) is recorded as a function of the scattering vector q ($q=4\pi\sin(\theta)/\lambda$, where 2θ equals the scattering angle). In SAXS analysis the solvent scattering is subtracted from the scattering intensity of the solution. For monodisperse non-interacting particles, the random positions and orientations of particles result in an isotropic intensity distribution which is proportional to the scattering from a single particle averaged over all orientations. In the case of biologic samples, it is possible to extrapolate information about the shape and size of the macromolecules in solution from the angular range, the scattering curves and the intensities. It is also possible to determine protein oligomeric states, folding/unfolding and surface-to-volume ratios. The parameters commonly determined by SAXS are listed in **Table 3** (for review see [126]).

Technically, the method requires only a minimum sample preparation; it is accurate and mostly non-destructive (since macromolecules are in solution). The experiments are fast and simple: they consist in injecting solutions of macromolecules into a capillary (the sample chamber), exposing them to X-ray and recording the scattered radiation **(Figure 28)**. Radiation damage is minimized by continuously flowing the solution in the capillary during X-ray exposure. SAXS experiments are often performed at synchrotron radiation sources that provide high intensity X-ray photon beams. Since SAXS measurements are performed very close to the primary beam (small angles), they require a highly collimated or focused X-ray beam. Moreover, macromolecules such as proteins usually scatter weakly and the solutions measured by SAXS are diluted.

SAXS has some advantages in comparison with other structure determination methods, such as X-ray crystallography and nuclear magnetic resonance (NMR). Unlike X-ray crystallography, SAXS does not require a crystalline sample for measurements and allows for fast protein structure determination (even if it is not possible to get to atomic resolution, range 10-50Å). An advantage of SAXS over NMR is that SAXS measurements are not limited by macromolecules sizes, whereas NMR is encountering problems with macromolecules of molecular masses over 30 kDa

Table 3. Parameters commonly determined by SAXS and their significance.

.

Parameters	Significance
R_g	Radius of gyration. Describes the mass distribution of the macromolecule around its centre of gravity. It can be calculated from Guinier or Debye Approximations and from p(r).
I(0)	Intensity at q=0 (or extrapolated intensity at zero scattering angle). It is also proportional to M and V.
p(r)	Pair distribution function. Describes the paired-set of distances between all of the electrons within the macromolecular structure.
D_{max}	Maximum dimension of particle. It is the maximum distance present in the scattering particle.
M	Mass. It is possible to calculate the mass from the I(0) expressed in absolute scale and normalized by mass/volume and not molar concentration.
V	Particle Volume. The volume of the macromolecule undergoing scattering. It can be calculated from I(0) and from the Porod invariant.

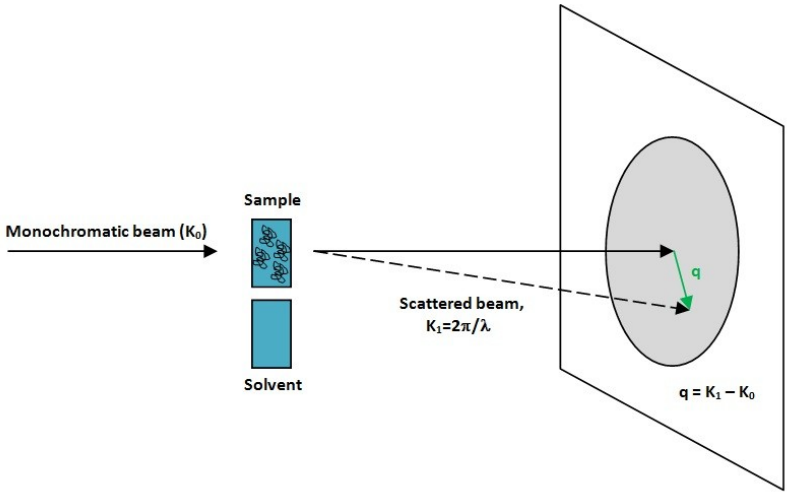


Figure 28 **Standard scheme of a SAXS experiment.**

2.9. CRYSTALLIZATION TRIALS AND X-RAY DATA COLLECTION

2.9.1. Crystallization of the p300 'core' and data collection

In order to obtain homogeneous, stable and crystallizable preparation of the p300 'core', purified untagged protein from *Hi5* cells was treated with the SIRT2 deacetylase (ratio 1:10, SIRT2:p300) overnight at RT. The reaction conditions were the same as those used in deacetylation assays. SIRT2 was separated from the p300 deacetylated 'core' by size exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare) equilibrated in buffer 20 mM Hepes pH7.5, 300 mM NaCl, 5 μM $ZnCl_2$ and 0.5 mM TCEP. The deacetylated protein was then concentrated to 8 mg/ml and incubated on ice for 30 minutes with a 3-fold molar excess of the inhibitor LysCoA, which was previously reported to stabilize the HAT domain [92]. The deacetylated LysCoA-p300 complex was finally centrifuged at 13000 rpm for 15 minutes and set up for crystallization.

Crystals of the p300 'core' were obtained with an initial screen of 576 different crystallization conditions in a 96-well sitting drop format using 100 nl drops. Crystals grown at 20°C by the vapour diffusion method emerged after 24 hours under multiple conditions. Crystallization conditions were then optimized manually by hanging drop vapour diffusion and with automated custom screens. Crystals were observed under many conditions ranging from pH 7.0-8.0, in the presence of several salts (0.2 M) and polyethylene glycols (PEGs) with molecular weight from 2000 to 8000 (16-22%). Different

protein concentrations were also tested (8 mg/ml to 15 mg/ml) and co-crystallization trials with BP substrates were performed. The best-diffracting crystals were obtained in hanging drop under two conditions: PEG 6000 (16-20%), pH 7.0-8.0 and PEG 3350 (18-22%), 0.2 M salt.

Data collection was performed at the ID14-4 beamline (ESRF, Grenoble). The majority of p300 crystals diffracted very anisotropically. In order to overcome this problem, several techniques were attempted such as changing of cryogenic conditions, annealing and dehydration. Best diffraction patterns were obtained from crystals frozen in 20-25% ethylene glycol or dehydrated by equilibration against precipitant solution at higher PEG concentrations (30-35%). Diffraction data on these crystals were collected using the EDNA strategy program as a guide [128].

2.9.2. p300 'core' crystal structure determination and structural analysis

The intensities obtained from the best diffracting crystal were integrated and scaled using the program XDS [129]. The crystal belonged to the space group C2 and diffracted at a maximum Bragg spacing of 2.8 Å. The structure was determined by molecular replacement with the program Phaser (CCP4i suite) using the p300 HAT domain as search models (Protein Data Bank (PDB) entry 3BIY). The BRD (PDB: ID 3I3J) was placed by FFFEAR (CCP4i suite).

2.9.3. BP crystallization

The p300 BP crystals were first obtained with untagged protein at 4°C in many conditions ranging from pH 6.0-7.0 and ammonium sulphate at different concentrations (1.5-1.7 M). Crystals grew as clusters of needles in all conditions. Crystal optimization was performed by hand and with a robot using custom screens. Several techniques were attempted to improve these crystals, including hanging drop, sitting drop, additive screen (robot/manual), *in situ* proteolysis, microseeding and macroseeding. Crystallization conditions were also varied. Broad pH ranges were screened (pH 4.0-pH 7.5) as well as ammonium sulphate concentrations (0.8-1.9 M). Different temperatures (4°C and 20°C) and protein concentrations (7 mg/ml, 10 mg/ml, 15 mg/ml, 30 mg/ml) were also tested as well as co-crystallization with substrate peptides. Crystals were tested for diffraction at the ID23-2 microfocus beamline (ESRF). It was not possible so far to obtain well diffracting crystals.

The CBP p300 BP crystals were also obtained at 4°C. These crystals where obtained under a single condition: 0.1 M Tris/HCl pH 8.5, 0.2 M NaCl and 25% PEG 3350. Crystals were reproduced manually under the same condition by hanging drop. Crystals were tested at the ID14-4 beamline (ESRF) but it was not possible to obtain good diffraction data.

2.9.4. What is X-ray crystallography?

X-ray crystallography is a method to determine the arrangement of atoms within a crystal. This experimental technique exploits the fact that X-rays are diffracted by crystals. It was originally developed with small molecules such as salts, minerals and various organic compounds. In the past years, it has become increasingly important in biochemistry given the discovery that proteins can be crystallized and that these crystals diffract in a similar way as small molecules.

The bottleneck in protein X-ray crystallography is to obtain good quality crystals for diffraction measurements. Once such crystals are obtained, they are transferred into cryogenic solutions, fished through nylon loops and flash-cooled in liquid N_2. This is done to slow down radiation damage upon crystal exposure to the X-ray beam. In a diffraction experiment, the flash frozen crystal is mounted on a goniometer and constantly mantained under a flow of N_2 stream (100 K). The crystal is gradually rotated while being shot with X-

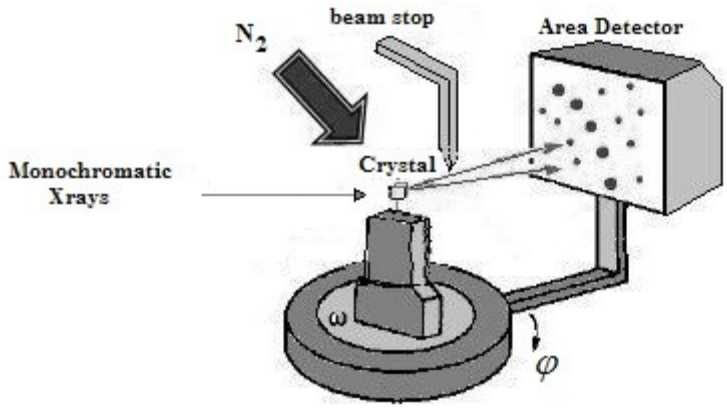


Figure 29 **Scheme of an X-ray experiment.** The goniometer allows for crystal rotation of an angle $\Delta\varphi$ along an axes perpendicular to the incident beam.

rays producing a diffraction pattern of regularly spaced spots (reflections). The signal due to the direct X-ray beam that passes through the crystal is stopped by a thin beam-stop. The diffracted X-rays are collected on an area detector and sent directly to computers for analysis **(Figure 29)**. Typically, a complete dataset from an X-ray diffraction experiment on a protein crystal consists of 50-1000 diffraction images.

Crystals are regular arrays of atoms and X-rays can be considered waves of electromagnetic radiation. X-rays have the appropriate wavelength (in the Å range, $\sim 10^{-8}$cm) to resolve the electron cloud of neighbouring atoms. Any wave colliding with a regular array of scatterers produces a regular array of secondary spherical waves. These waves are cancelled in most directions through destructive interference. However, they add constructively in a few specific directions, determined by Bragg's law:

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

where d is the spacing between diffracting planes, θ is the incident angle, n is any integer, and λ is the wavelength of the incident beam. These specific directions appear as spots (or reflections) on the diffraction pattern.

The diffraction pattern obtained is used to determine cell parameters and the symmetry of the crystal, or space group (indexing) [129, 130]. In this way, it is possible to obtain a record of the Miller Index (*h*,*k*,*l*) of each reflection. Having assigned the symmetry of the crystal, the integration of the data allows to associate an intensity to each measured *hkl* reflection [130, 131]. These integrated intensities are affected by scaling problems since they can change from one recorded image to another. Indeed, in a complete dataset hundreds of separate images are recorded at different orientation of the crystal. Optimizing the intensity scale (scaling) is critical because the relative intensity of the peaks is the key information from which the structure is determined [130, 132]. Indeed, the intensity is proportional to the square of the structure factor amplitude:

$$I_{hkl} = |F_{hkl}|^2$$

where I_{hkl} and F_{hkl} are the intensity and the structure factor of a reflection characterized by the Miller index *hkl*, respectively. The structure factor is the description of how the atoms in the molecules within the crystal scatter the incident radiation. It is a complex number containing information on both amplitude and phase of the wave. Therefore, in order to obtain an interpretable electron density map, both amplitude and phase must be known.

For measurement on protein crystals, the phase cannot be recorded directly during the diffraction experiment. Phase information can be obtained by performing additional experiments on crystals soaked with solutions containing strong scatterers such as heavy atom solutions or by collecting anomalous data on protein derivatives (e.g. SeMet).

However, when the structure of a protein related to the one currently studied is available (sharing a predicted identical fold and sequence identity higher than 25-30%), a model from the homolog structure can be used to determine the orientation and the position of the molecule within the cell. This technique is called molecular replacement [133]. The phases obtained allow to calculate an initial electron density map. A model is then progressively built into the electron density map and refined against the experimental data. For detailed explanation of the protein crystallography pipeline see I. Drent [134] and B. Rupp [135].

2.10. COLORIMETRIC ACTIVITY ASSAYS

A colorimetric HAT assay was developed based on that by Thompson *et al.* [6] in order to test the influence of the BP module on HAT activity. This assay uses the Ellman's reagent or DTNB (5,5'-dithio-bis-(2-nitrobenzoic acid)) to detect the production of CoASH (coenzyme A) molecules during the acetylation of a histone substrate. DTNB can react with SH groups and produce a chromophore (thiophenolate) which absorbs at 412 nm. In this particular case, CoASH reacts with DTNB producing equimolar amount of thiophenolate. Acetylation can then be simply quantified (production of one CoASH molecule equals one

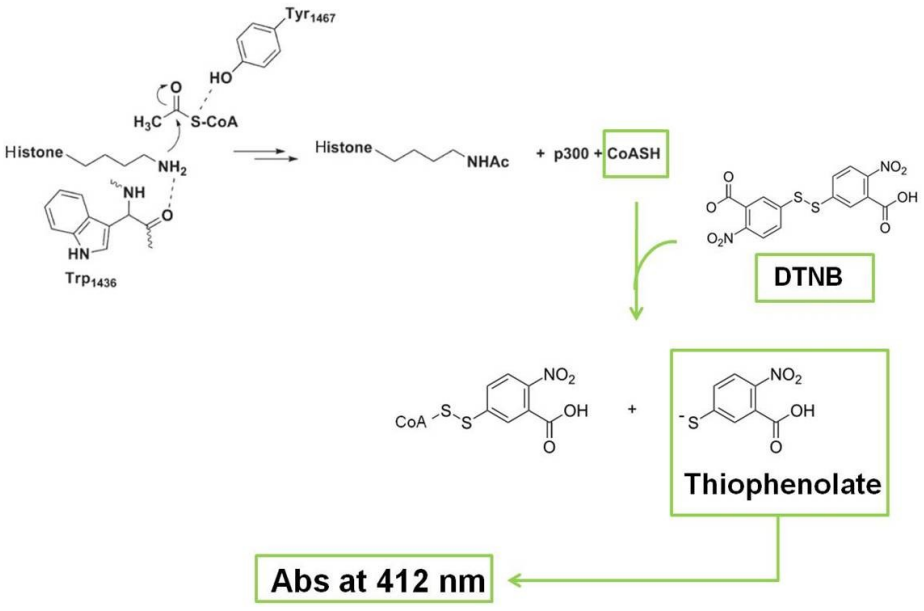


Figure 30 **The DTNB colorimetric assay.** This coupled assay allows for the quantification of peptide acetylation by monitoring the generation of thiophenolate production (abs 412 nm) from the reaction of CoASH with DTNB.

acetylation event) by measuring the absorbance at 412 nm **(Figure 30)**.

For the DTNB assays, reactions were performed in 30 µl of 100 mM sodium potassium phosphate pH7.5, 300 mM NaCl, 4 mM AcCoA. The HAT Δloop construct (2 µM final, putative constitutively active protein) was incubated with 0-20-40 µM of the BP module. Concentration of BP was mantained constant for each experiment whereas the concentration of the substrate peptide H3 (20-mer) was varied from 0 mM to 2 mM to measure k_{cat} and K_m parameters. In preliminary assays, it was determined that enzyme activity was in the linear range for about 30 minutes at 30°C and that the concentration of AcCoA used in the experiment was not a limiting factor in the acetylation reaction.

The reaction in presence of BP was carried out at 30°C for 15 minutes. Enzyme activity was quenched by addition of 30 µl guanidinium hydrochloride (6 M). In order to measure CoASH production, 30 µl of DTNB (6 mM) were added to the quenched reactions and the absorbance at 412 nm was recorded using a Victor 1420 multilabel counter (Wallac). Assays were performed at least in triplicates. Background absorbance was determined and subtracted from the absorbance determined for individual reactions. The absorbance values were then converted into CoASH concentration using the mean conversion factor calculated from a standardization curve done using serial dilutions of CoASH (Sigma).

The apparent k_{cat} and K_m for the H3 peptide substrate in presence of fixed concentrations of BP were determined from the intercepts of the double reciprocal (or Lineweaver–Burk) plot. Equations are reported in **Figure 31**.

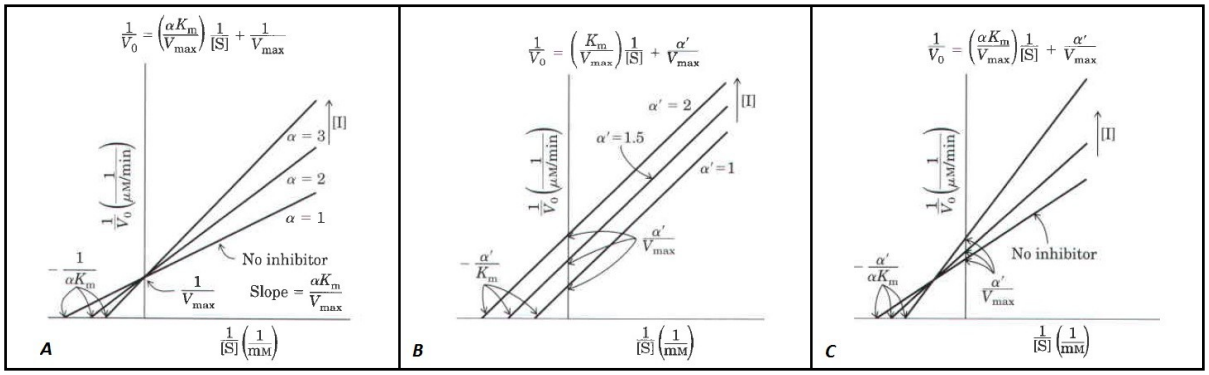


Figure 31 **Sets of double reciprocal plots.** Plots are obtained in absence or presence of inhibitor at two different concentrations [1]. (A) Competitive inhibition. (B) Uncompetitive inhibition. (C) Mixed inhibition.

3. RESULTS

RÉSUMÉ EN FRANÇAIS

L'expression et la purification du 'noyau' de p300 ont été les plus grands défis dans la caractérisation de cette enzyme. Il a été possible d'obtenir de petites quantités de la protéine native (*wild type)* pour tester l'activité enzymatique sur des nucléosomes reconstitués et sur des histones recombinantes. Des quantités plus importantes de cette protéine ont été obtenues après expression dans des cellules d'insecte *Hi5* et grâce à l'introduction d'une mutation (Y1467F) dans le site catalytique de p300 qui remplace un résidu de tyrosine par une phénylalanine. Les protéines contenant cette mutation ont une activité catalytique réduite par rapport à la protéine native. En plus, une délétion de la boucle d'auto-inhibition comprenant les résidus 1520-1580 a été introduite afin d'améliorer le rendement de l'expression de la protéine. La p300 Δ1520-1580/Y1467F protéine est hypo-acétylée de façon hétérogène suite à l'expression hétérologue. Les tests de protéolyse réalisés sur les différentes constructions de p300 acétylées différemment avancent la possibilité que certains changements de conformation peuvent avoir lieu à l'intérieur de la protéine au cours des transitions entre les différents états d'acétylation.

Grâce au traitement de p300 Δ1520-1580/Y1467F avec une désacétylase appelée SIRT2, il a été possible d'obtenir une protéine homogène et desacétyle. La protéine ainsi préparée a permis de produire des cristaux de bonne qualité pour la caractérisation structurale par rayon X. Un model partiel à haute résolution (2.8 Å) a été obtenu par le traitement des données recueillies pour un de ces cristaux sur la ligne ID14-4 (ESRF, Grenoble). Ce modèle a été confirmé par les données de diffusion des rayons X aux petits angles (SAXS) collectées sur la ligne SWING (SOLEIL, Paris).

Le module BP, le domaine HAT et la désacétylase SIRT2 ont été exprimés dans *E. coli* BL21 (DE3) et purifiés à l'homogénéité: ce sont tous des monomères en solution. Dans le cas du domaine HAT, la délétion Δ1520-1580 et la mutation Y1467F ont été introduites afin d'améliorer le rendement de l'expression de la protéine. Dans le cas du module BP, des mutants ponctuels dans des résidus clés du bromodomaine et du PHD *finger*, les mutations N1131A et K1253A respectivement, ont été préparés.

Des études d'interaction entre le module BP et le domaine HAT ont été effectuées en utilisant trois techniques principales: la calorimétrie de titration isotherme (ITC), GST-*pull*-

downs et tests par chromatographie d'exclusion stérique. Ces études montrent que BP peut former une interaction stable avec le domaine HAT et que cette interaction nécessite une surface de liaison différente de celle habituellement utilisée par le bromodomaine et le PHD *finger* dans la reconnaissance des substrats. En outre, des mesures d'activité enzymatique ont été effectuées par dosages colorimétriques afin de déterminer l'influence du module BP sur l'activité du domaine HAT. Les données ont démontré que le module BP a une action inhibitrice sur l'activité du domaine HAT.

Enfin, des essais d'interaction ont été effectués pour déterminer la spécificité de substrat du module BP. Le module BP peut reconnaître des lysines acétylées sur des peptides d'histone en discriminant entres les mono- et di-acétylations. Ces lysines ainsi identifiées sont les mêmes résidus qui sont modifiés par p300 *in vivo* et *in vitro*. En conclusion, les données obtenues ont permis de déterminer une des caractéristiques requises par le module BP afin de se lier spécifiquement à ses substrats: la présence de trois résidus, préférentiellement petits, entre les deux lysines acétylées reconnues par le module.

3. RESULTS

3.1. PROTEIN EXPRESSION AND PURIFICATION

3.1.1. The p300/CBP 'core'

The central region ('core') of p300 (aa 1048-1666, 75kDa), comprising the BP module and the HAT domain, was recombinantly expressed in *Hi5* and *E.coli* BL21 (DE3). It was purified to homogeneity, though at very low yields (0.13 mg/L from both hosts). It was possible to demonstrate by HAT assays that the recombinant wild type protein was active on reconstituted nucleosomes and on full length free histones (H3 and H4 from *D. melanogaster*). A mutant of the p300 core bearing Y1467F (Tyr>Phe) mutation was also generated. Tyr 1467 is a crucial residue for p300-mediated acetyl-transfer catalysis. It was shown that this Y1467F mutant possessed reduced HAT activity **(Figure 32)**.

Mass spectrometry (MS) analysis of the p300 'core' constructs revealed that the p300 *wt* was hyperacetylated whereas the Y1467F mutant was hypoacetylated upon

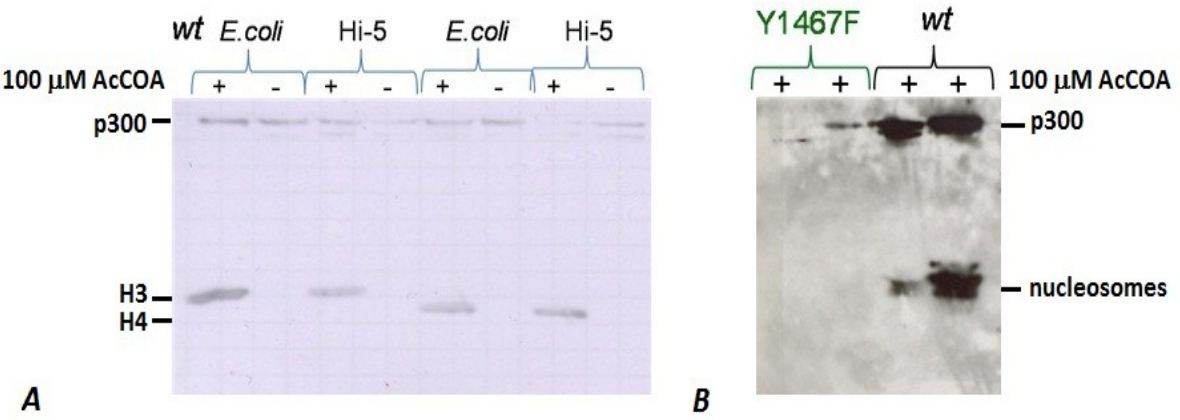


Figure 32 **HAT assays on reconstituted nucleosomes and free full length histones.** Levels of acetylation were detected by western blot using an anti-acetyllysine antibody. (A) For the free full-length histones, activity was measured in presence or absence of AcCoA using the p300 *wt* constructs from both *E.coli* BL21 (DE3) and *Hi5*. The p300 constructs were autoacetylated upon overexpression in host cells. In absence of AcCoA, the p300 constructs were not active on free full-length histones. (B) Activity on reconstituted nucleosomes was measured using p300 *wt* and Y1467F mutant from *Hi5*. The p300 Y1467F mutant was hypoacetylated if compared with the p300 *wt* and it was not active on reconstituted nucleosomes.

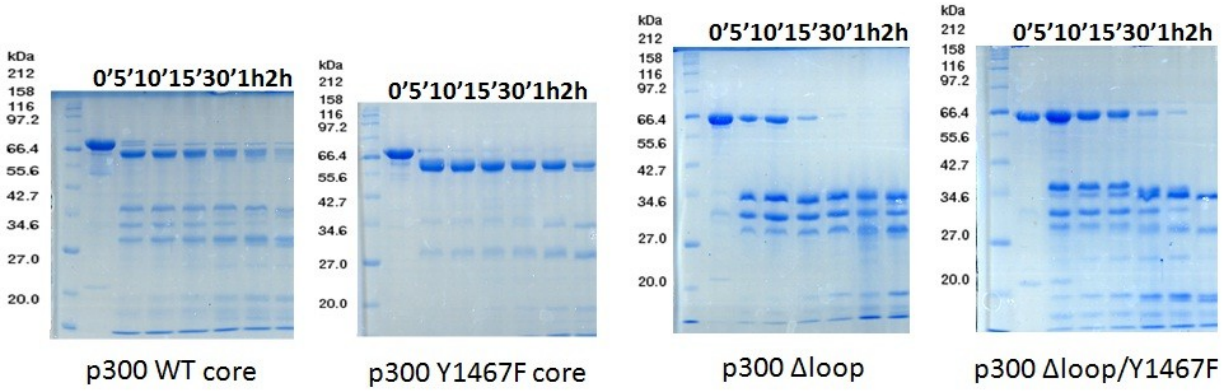


Figure 33 **SDS-page analysis of the trypsin cleavage of p300 constructs.** Incubation time in minutes (') and hours (h) is indicated above the gels. The names of the proteins as well as the molecular weight (kDa) of the marker are given. Protein fragments were visualized by Coomassie blue staining.

overexpression both in *Hi5* and in *E.coli* BL21 (DE3). I probed the structural consequences of auto-acetylation by limited proteolysis. The p300 'core' *wt* and Y1467F mutant constructs exhibited slightly different banding patterns and cleavage kinetics suggesting that a conformational change may occur upon autoacetlyation **(Figure 33)**. As previously suggested by Liu *et al.* [92] these differences might be due to the displacement of the autoinhibitory loop, which might be folded back to mask the active site and stabilize the autoinhibited conformation of p300, when hypoacetylated.

In order to improve protein yields suitable for crystallization, new p300 constructs were designed. Inspection of the structure of the p300 HAT domain suggested that it was possible to delete the autoinhibitory loop and replace it with a small linker (Δ1520-1580 or Δloop) **(Figure 34)**. This loop contains many lysine residues and is very flexible, which may represent a problem for protein expression and for the subsequent crystallization.

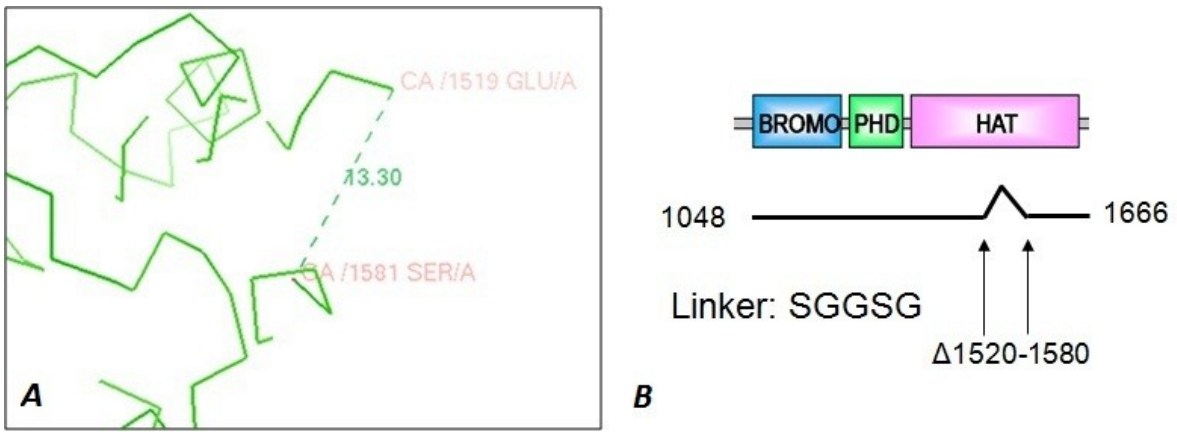


Figure 34 **Scheme of p300 loop region deletion.** (A) The p300 autoinhibitory loop was deleted in the atomic structure available for the p300 HAT domain (PDB: ID 3BIY). Here is represented a zoom-in of the region containing the N- (Glu1519) and C- (1581) termini of the autoinhibitory loop. The distance between the N- and C-termini is given in Å. (B) Scheme of the p300 construct cloned. The linker connecting the N- and C- termini of the autoinhibitory loop is given.

Previously, for crystallization of the HAT domain, this loop was proteolytically cleaved off [92]. Both p300 Δ1520-1580 and Δ1520-1580/Y1467F constructs were poorly expressed in *E.coli* (0.3 mg/L). Limited proteolysis on both Δloop constructs gave different banding patterns and cleavage kinetics as compared to the full length 'core' constructs. Overall, loop deletion allowed design of a correctly folded protein and autoacetylation resulted in small conformational differences **(Figure 33)**.

The p300 Δ1520-1580/Y1467F construct was expressed at good yields in *Hi5* (5-6 mg/L) and purified to homogeneity. In contrast, expression of p300Δloop resulted in very low yields presumably due to toxicity of the *wt* protein for the host upon overexpression (constructs lacking the autoinhibitory loop are constitutively active [6]). The p300 Δ1520-1580/Y1467F protein was characterized by MS, Multi-Angle Laser Light Scattering (MALLS) and Dynamic Light Scattering (DLS). It was heterogeneously hypoacetylated upon overexpression in insect cells **(Figure 35)**, as revealed by mass spectrometry analysis: up to 5 to 6 acetylation marks instead of the 13 to 17 for the *wt* protein.

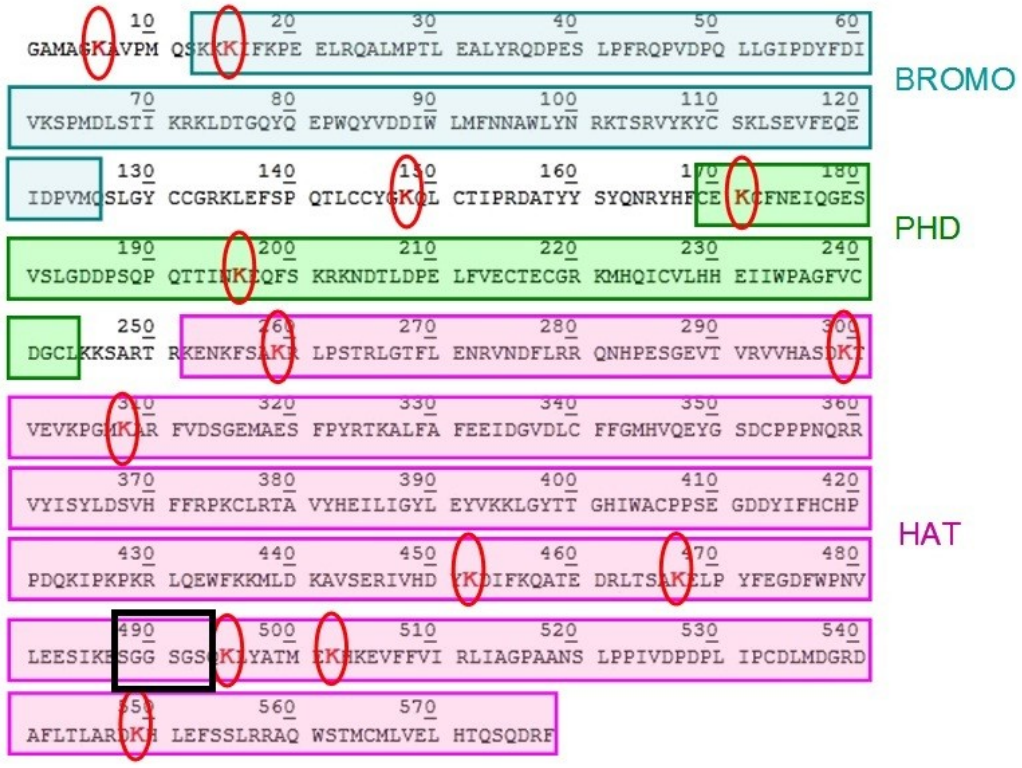


Figure 35 **Acetylated residues on p300 'core' Δ1520-1580/Y1467F.** The picture shows the sequence of the p300 'core'. In red circles are indicated the acetylated residues identified by mass spectrometry analysis. In blue, green and pink boxes are indicated the regions of the sequence corresponding to the bromodomain, the PHD finger and the HAT domain, respectively. In the black box is shown the linker (SGGSG) between the N- and C- termini of the autoinhibitory loop. Overall, each p300 molecule presented a combination of 5 to 6 acetylated residues among the possible ones.

3.1.2. BP module

The p300 bromo-PHD *wt* was expressed in *E.coli* BL21 (DE3) and purified to homogeneity at good yields (2-4 mg/L) **(Figure 36)**. An initial 1D NMR spectrum indicated that the BP construct is well-folded and globular (data not shown). Limited proteolysis studies showed that the BP *wt* was resistant to trypsin digestion, suggesting that the bromo- and the PHD- domains could form a structural unit.

Two BP mutants were designed to study the function of the BP module. In the N1131A mutant, a conserved asparagine implicated in AcLys binding was replaced by an alanine. In the K1253A mutant, a lysine thought to be important for the recognition of methyl-lysines was mutated into alanine. These constructs were expressed in *E.coli* BL21 (DE3) and purified to homogeneity with yields comparable to the *wt* BP (2-4 mg/L). They were globular monomeric species as revealed by size exclusion chromatography.

The CBP BP construct was also expressed in *E.coli* BL21 (DE3) and purified at good yields (2-3 mg/L) as a globular monomeric species.

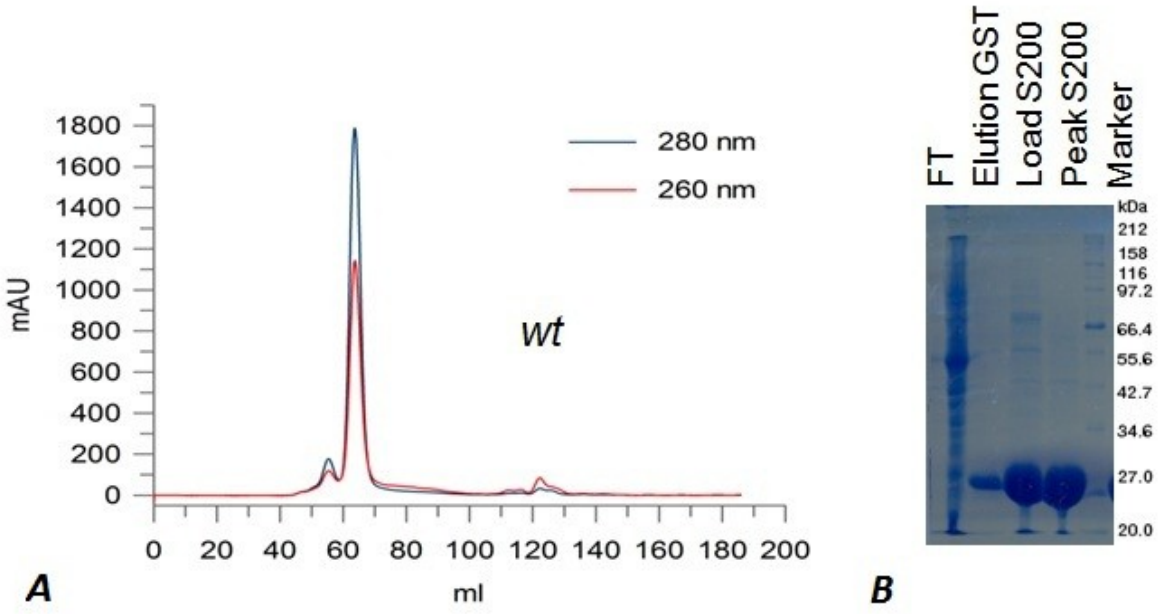


Figure 36 **Purification of p300 BP *wt*.** (A) Chromatogram showing the elution profile from the Superdex S75 gel filtration column. Absorbances at 260 nm and 280 nm were measured to detect protein elution. (B) SDS-page gel showing the protein samples taken after each purification step. Purification steps are indicated above the gel. Molecular weight (kDa) of the marker is given. Proteins were visualized by Coomassie blue staining.

3.1.2. The HAT domain and the deacetylase SIRT2

The p300 HAT domain *wt* and Y1467F constructs were poorly expressed in *E.coli* BL21 (DE3) and BL21 (DE3/Rosetta) (0.2 mg/L). In order to improve protein yields, constructs with deletion in the autoinhibitory loop region were designed in a similar way as for the p300 'core'. HAT Δ1520-1580 and HAT Δ1520-1580/Y1467F could be expressed at good yields in *E.coli* (1 mg/L and 8-9 mg/L respectively) and purified to homogeneity.

In order to perform deacetylation assays, the SIRT2 deacetylase was expressed at good yields in *Hi5* cells (20 mg/L) and purified to homogeneity. Deacetylation by SIRT2 was carried out on the p300 'core' Δ1520-1580/Y1467F construct **(Figure 37)**. Western blot using anti-acetyllysine antibody and mass spectrometry analysis were performed in order to detect acetylation levels.

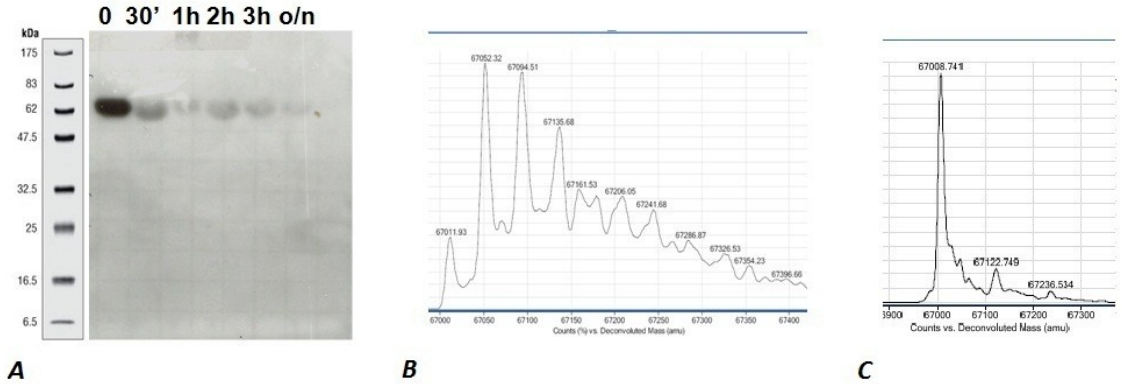


Figure 37 **Western blot and mass spectrometry analysis of p300 deacetylation by SIRT2.** (A) Levels of acetylation were detected by western blot using an anti-acetyllysine antibody. Time in minutes (') and hours (h) is indicated above the gel. The molecular weight (kDa) of the marker is given. After overnight (o/n) reaction, SIRT2 was able to completely deacetylate the p300 'core'. (B-C) Mass spectrometry analysis using electrospray ionization (ESI) of the p300 'core' before (B) and after (C) deacetylation by SIRT2. The p300 'core' has a calculated molecular weight of 67 kDa. Each acetylation event adds 42 Da.

3.2. INTERACTION BETWEEN THE BP MODULE AND THE HAT DOMAIN

3.2.1. BP-HAT binding studies

The close sequence proximity of the BP module and the HAT domain suggests that these two domains perform concerted functions and might physically interact during their reaction cycle. We speculated that the BP might 'sense' the HAT activation status by binding to the HAT domain and/or the autoinhibitory loop. For instance, the BP-HAT interaction may stabilize the inactive autoinhibited conformation of the p300 'core'. In order to test this hypothesis, ITC studies were performed using untagged BP module and untagged HAT Δloop/Y1467F. BP mutants in the BRD (N1131A) and in the PHD domain (K1253A) were also tested for binding to the HAT domain. ITC data showed that the BP module binds tightly to the HAT domain and that this interaction is not much affected by mutations within the bromo- and PHD- recognition sites for PTMs **(Figure 38)**. Measured ITC parameters for each one of the BP constructs are reported in **table 4.**

Next, the BP-HAT interaction was confirmed by analytical size exclusion chromatography, which is a useful tool to detect complex formation between two proteins. Formation of the BP-HAT complex resulted in a higher molecular weight species, characterized by an elution profile significantly shifted if compared to the individual free proteins. HAT and BP were mixed in equimolar ratio for the experiment **(Figure 39)**.

Table 4. Thermodynamic analysis of the interaction between BP and the HAT domain. The variability in the experimental values were calculated to be about 1% in the number of binding sites (*n*), 5% in the binding enthalpy (ΔH) and 5% in the dissociation constant (K_d).

BP construct	*n*	ΔH (kcal/mol^{-1})	K_d (μM)
wt	0.73	-1.502 e^{-4}	1.67
N1131A	1.09	-1.785 e^{-4}	2.3
K1253A	1.08	-2.646 e^{-4}	18.2

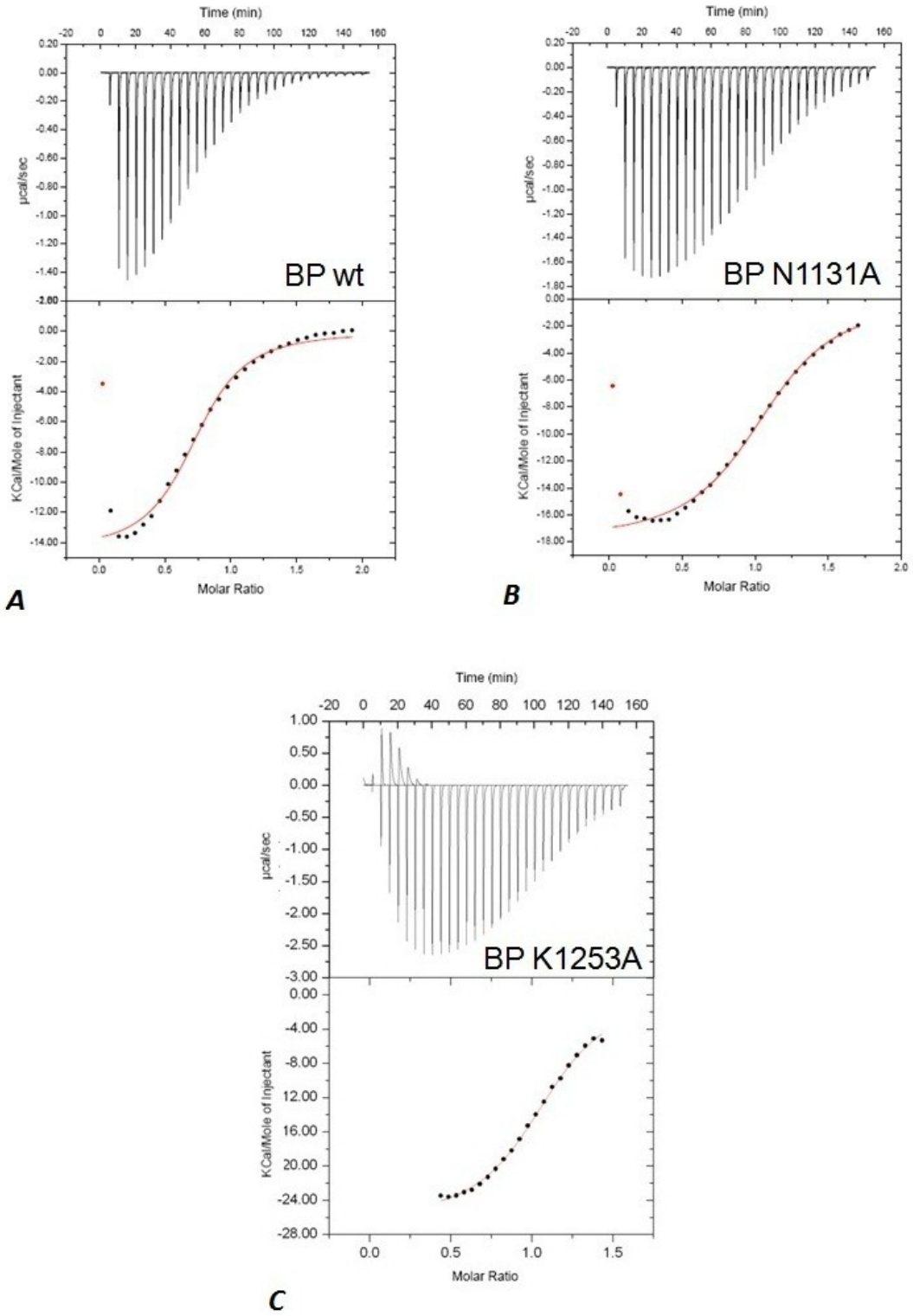


Figure 38 **ITC and pull-downs assays to test BP-HAT interaction.** Calorimetric titration of the p300 HAT domain with the BP module. Experiments were performed using VP-ITC and constant volume of injection (7µl). Black dots represent the experimental data, continuous red lines correspond to the best fit to a model considering one set of binding sites.

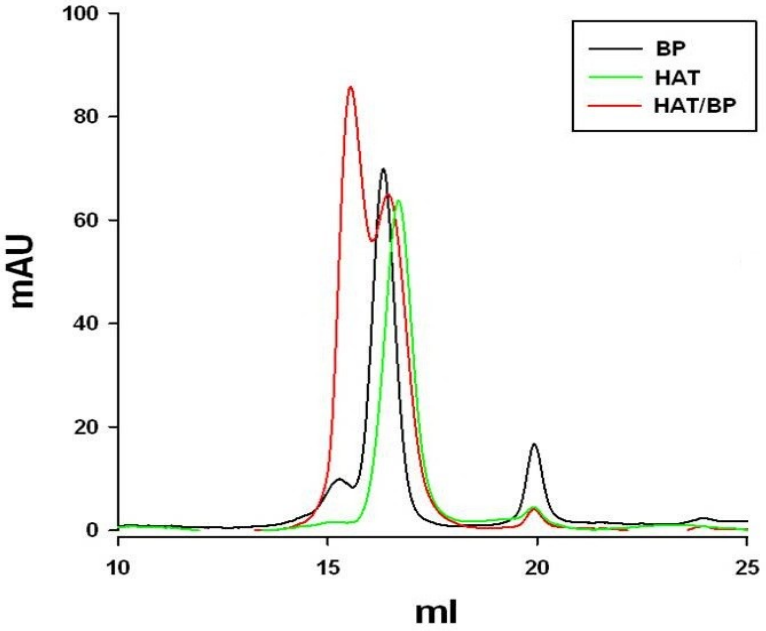


Figure 39 **Chromatograms of size exclusion analytical runs.** Chromatograms showing the elution profiles of the BP module (black), the HAT domain (green) and the complex BP-HAT (red) using a Superdex S75 gel filtration column. Absorbance at 280 nm was measured to detect protein elution.

The BP-HAT interaction was further confirmed by GST-pull-down assays using GST-tagged BP **(Figure 45)**. The addition of a BP substrate (H3K14K18 diacetylated peptide) in GST-pull-down assays did not impair the BP-HAT complex formation **(Figure 40)**.

Moreover, additional GST-pull down assays showed that BP *wt* was able to bind both hyperacetylated and deacetylated HAT, suggesting that the BP-HAT interaction is not influenced by the presence of different acetylation states on the HAT domain **(Figure 41)**.

Figure 40 **GST-pull-down using GST-BP and HAT Δloop/Y1467F.** SDS-PAGE gel showing a pull-down in presence (+) or absence (-) of a BP substrate (histone H3 peptide K14K18 diacetylated). The molecular weight (kDa) of the marker is given. Negative controls of the GST-BP and of the HAT domain alone are also shown (names are indicated above the gel). M, marker; I, input; FT, flow through; E, elution. Proteins are visualized by Coomassie blue staining.

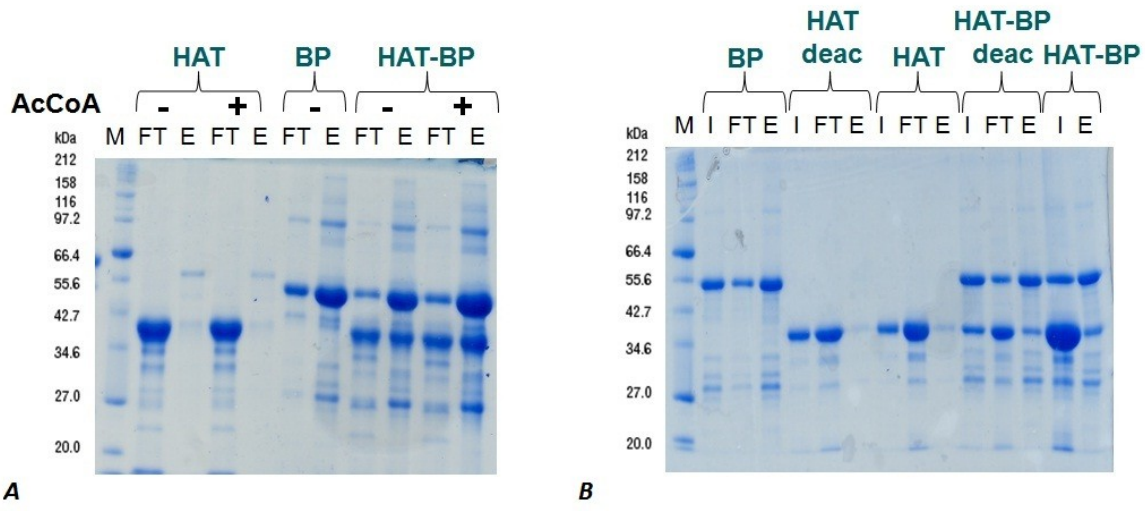


Figure 41 **GST pull-downs with HAT domains at different acetylation states.** SDS-PAGE gels showing GST-pull downs using hyper- (A) and de- acetylated (B) HAT Δ1520-1580 domain. Protein names are indicated above the gels. Negative controls of the HAT domain and the GST-BP alone are also shown. The molecular weight (kDa) of the marker is given. M, marker; I, input; FT, flow through; E, elution. Proteins are visualized by Coomassie blue staining.

3.2.2. HAT activity assays

HAT activity assays were performed in order to investigate the functional consequences of the interaction between the BP module and the p300 HAT domain. First, it was determined the linearity of the enzyme activity in a range of 30 minutes at 30°C **(Figure 42)**. Second, a calibration curve was calculated using serial dilution of CoASH. Kinetic studies were carried out for 15 minutes at 30°C using different concentrations of the BP module. Analysis of the double-reciprocal plots revealed a tendency of the BP module to

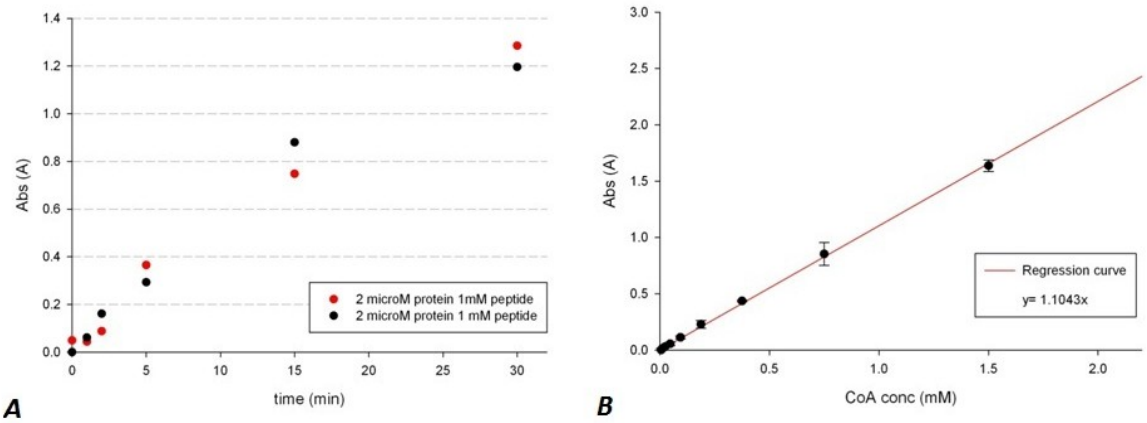
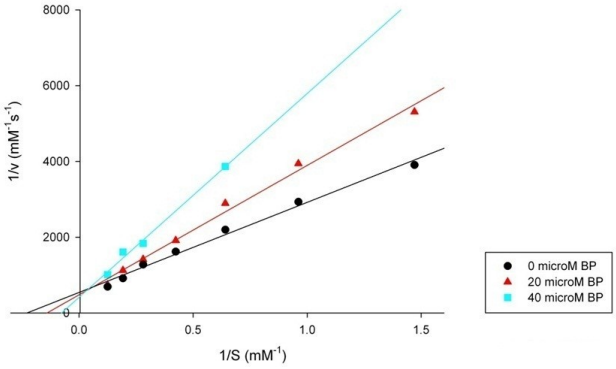


Figure 42 **HAT assay time course and CoASH standardization.** (A) A time course of the H3 peptide acetylation by p300 HAT Δ1520-1580 using a colorimetric DTNB coupled assay. Absorbance at 412 nm is plotted as a function of the time (min). The experiment was done in duplicate (red and black dots) at 30°C. (B) The standard curve calculated using serial dilution of CoASH. Absorbance at 412 nm due to the formation of thiophenolate from the DTNB reaction is plotted as a function of the CoASH concentration (mM). Experiments were carried out at 30°C and the reaction was stopped after 15 minutes.

act as a mixed inhibitor for p300 HAT acetyltransferase activity. As a mixed inhibitor, the BP module might bind to either the enzyme or the complex substrate-enzyme to affect the reaction. Apparent k_{cat} and K_m measured in these experiments are reported in **Figure 43**.



BP conc	K_m	K_{cat}
0 μM	4.39 ± 3.08 μM	0.92 ± 0.45 s^{-1}
20 μM	7.12 ± 3.1 μM	1.04 ± 0.44 s^{-1}
40 μM	13 ± 12.04 μM	1.2 ± 0.53 s^{-1}

Figure 43 **HAT assays.** Double-reciprocal plots of the initial velocities ($mM^{-1}s^{-1}$) versus substrate concentrations (mM^{-1}) are shown. Experiments were carried out at fixed concentrations of the BP module: 0 μM (black dots), 20 μM (red triangles), 40 μM (cyan squares). Substrate (histone H3 peptide) concentration was varied between 0 and 2 mM. In the table are reported the parameters obtained from the curves. Note that the calculated standard deviations are high due to variability of the DTNB reaction (which is temperature and humidity sensitive). The general trend of the double-reciprocal plots suggests a mechanism of mixed inhibition for the BP module (see Chapter 2.10).

3.3. STRUCTURAL STUDIES ON THE P300 'CORE'

3.3.1. Preliminary X-ray crystal structure

Parallel to the biochemical characterization, X-ray crystallographic studies were performed on the p300 'core'. A number of trials on different p300 'core' variants did not yield a crystallizable sample. Also, proteolysis of the core did not allow crystallization. Only after the construct p300 Δ1520-1580/Y1467F was deacetylated, I was able to obtain crystallizable material. The best diffracting crystals were obtained in hanging drops under two conditions: PEG 6000 (16-20%), pH 7.0-8.0 and PEG 3350 (18-22%), 0.2 M salt **(Figure 44)**.

Data on p300 Δ1520-1580/Y1467F crystals were collected on ID 14-4 (ESRF, Grenoble). The best-diffracting crystals diffracted to a minimum Bragg spacing of 2.8 Å. A partial molecular replacement solution was obtained by using the structure of the HAT domain (PDB: ID 3BIY) and the BRD of p300 (PDB: ID 3I3J) as search models. This was done using the programs Phaser and FFEAR. The latter allows for molecular replacement calculations when weak phases are available, such as in the case of the p300 'core'. Crystallographic statistics are reported in **Table 5**.

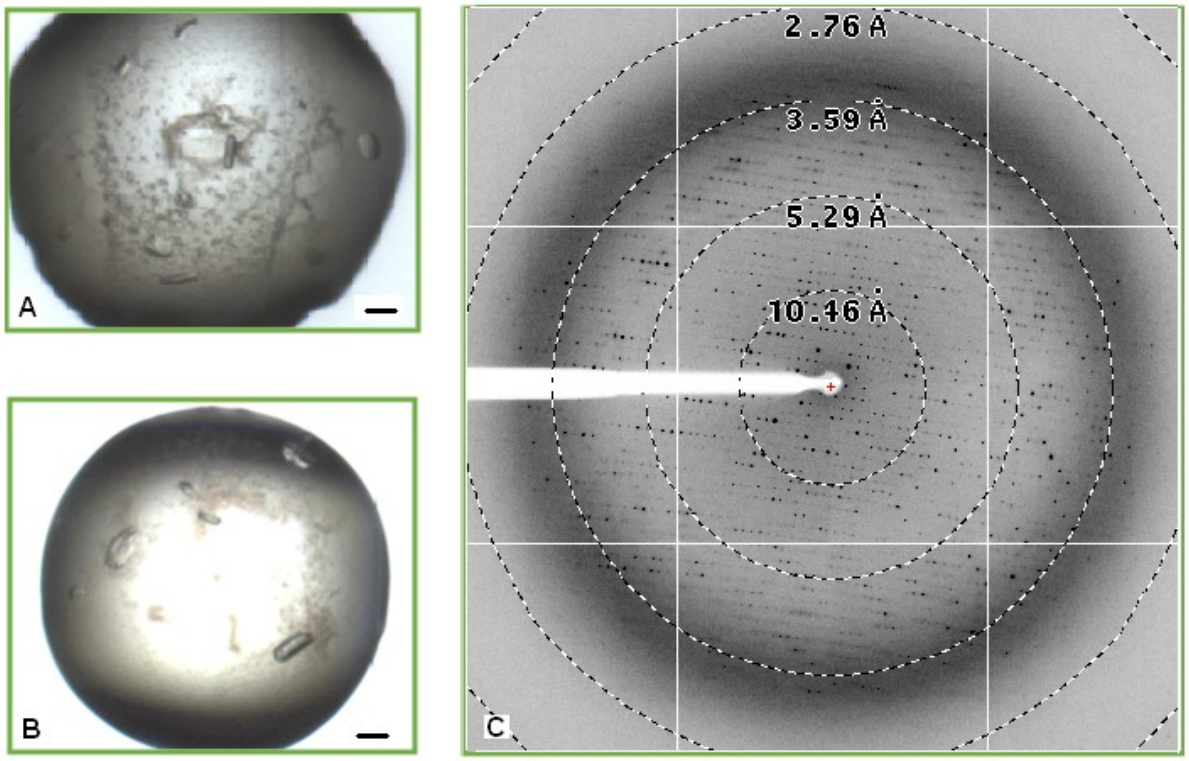


Figure 44 **p300 'core' crystals and diffraction pattern.** The p300 Δloop/Y1467F crystals obtained in PEG 6000 (A) and PEG 3350 (B) are shown. Black lines correspond to 0.1 mm. (C) Example of diffraction pattern obtained from the p300 'core' crystals. Images were collected on ID14-4 (ESRF).

Table 5. Data collection statistics. Values in parentheses are for highest-resolution shell.

Data Collection	Native
Space group	C2
Cell dimensions	
a, *b*, *c* (Å)	154.19, 92.48,118.43
β (°)	96.757
Wavelength	1.069
Resolution (Å)	50-2.80
R_{sym} or R_{merge}	14.2 (73.8)
$\langle I \rangle / \sigma I$	7.5 (1.7)
Completeness (%)	99.3 (99)
Redundancy	3.4 (3.4)

Preliminary characterization of the crystals by calculating the Matthews coefficient allowed determination of the solvent content. This analysis indicated two or three molecules per asymmetric unit **(Table 6)**.

In addition, a self-rotation function was calculated in order to determine if the macromolecules present within the asymmetric unit were related by rotational symmetry axes and if so, to determine the orientation of these axes within the unit cell. This allowed to detect presence of non-crystallographic symmetry (NCS). The analysis revealed the presence of a 2-fold non-crystallographic symmetry (Chi=180) **(Figure 45)**.

Information obtained from the self-rotation function calculation was in agreement with the data obtained from molecular replacement and with the calculation of the Matthews coefficient.

Table 6. Matthews calculations. Estimated molecular weight 67008 Da.

Nmol/asym	Matthews Coefficient	% solvent	P(3.50)	P(tot)
1	6.64	81.48	0.00	0.00
2	3.32	62.95	0.23	0.16
3	2.21	44.43	0.76	0.83
4	1.66	25.90	0.00	0.00
5	1.33	7.38	0.00	0.00

I conclude that the p300 'core' crystals belong to the monoclinic space group C2 and they are characterized by the presence of two macromolecules per asymmetric unit, which are related by a 2-fold NCS.

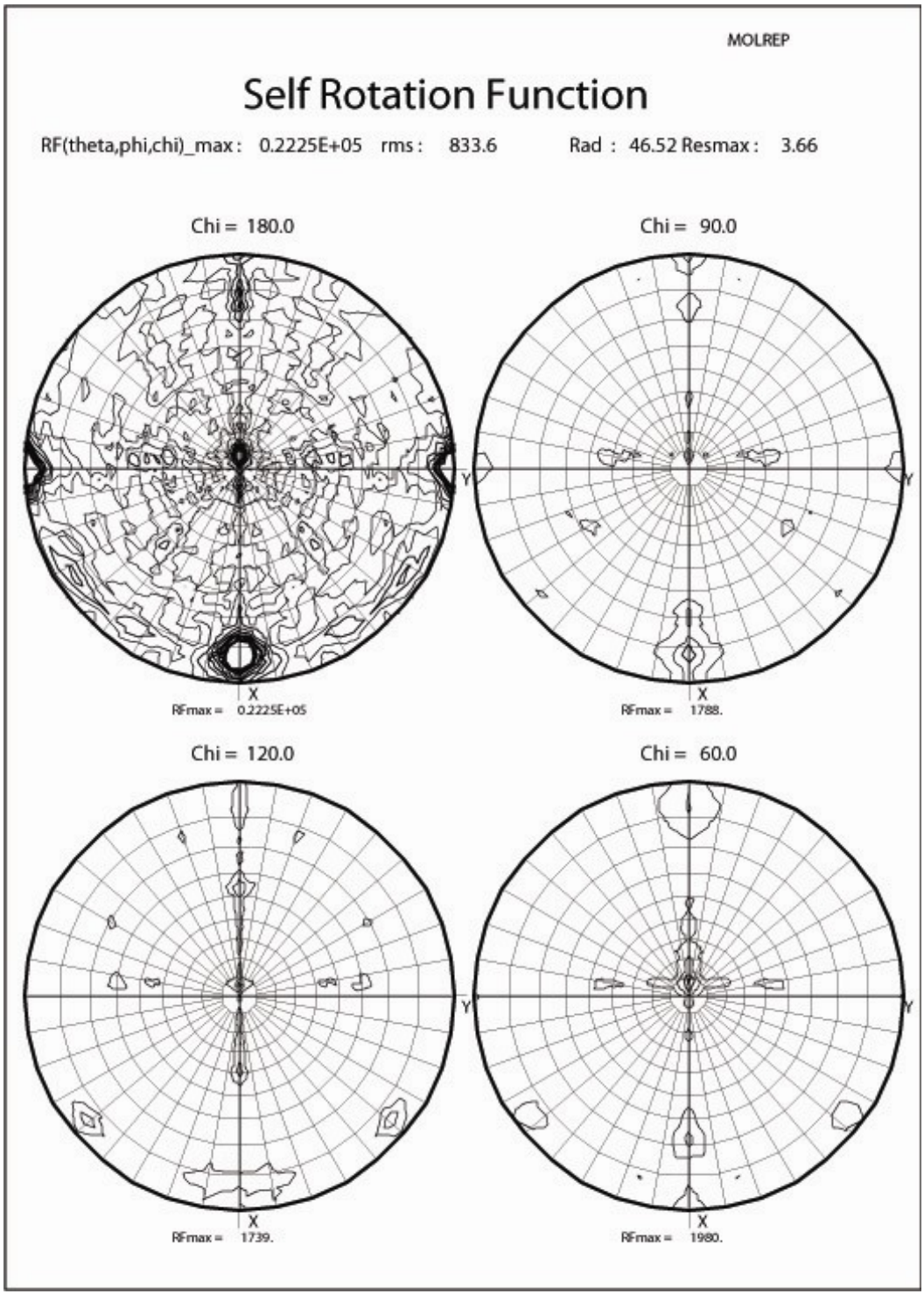


Figure 45 **Self rotation function calculated for the monoclinic C2 p300 'core' crystals.** Peaks in the Chi=180 section indicate a 2-fold non-crystallographic symmetry. Two macromolecules are present per asymmetric unit.

Attempts to solve the structure were performed as follows. First, molecular replacement was done with Phaser using the structure of the p300 HAT domain (PDB: ID 3BIY) and the bromodomain (PDB: ID 3I3J) as a search models. Two high scoring solutions for the HAT domain were found (LLG scores of +2556.45 and +2556.41). Map calculation showed good electron density for the HAT domains. Additional electron density was seen for the missing parts of the p300 'core', but the quality of the map was too poor for placement of the BRD or for model building. Next, I used FFFEAR to search for the BRD in the density map. I was able to find two solutions with reasonable scores. The best solution obtained is shown in **Figure 46**. However, even including the BRD, the quality of the map was too poor to allow for manual model building.

In order to locate the Zn^{2+} atoms in the PHD domain and improve phase information for the p300 'core', I measured data from the p300 'core' crystals at the Zn^{2+} ion absorption edge (wavelength 1.2837 Å). However, the anomalous signal obtained was too weak to allow calculation of anomalous difference maps.

Other techniques, such as cross-crystal averaging with density modification and NSC averaging have also been tried in order to improve phase information for the p300 'core' crystals [136]. However, so far these attempts did not lead to significantly better density

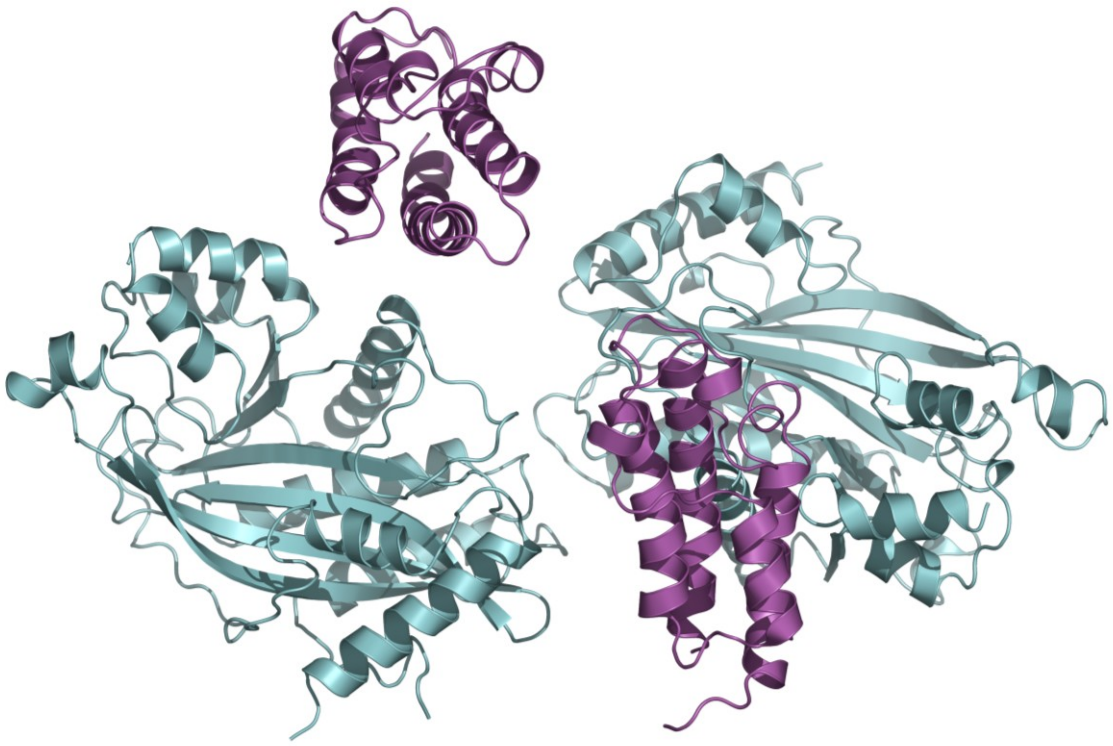


Figure 46 **Partial model of the p300 'core' showing the asymmetric unit.** The p300 'core' crystals are characterized by the presence of two macromolecules per asymmetric unit, related by a 2-fold NCS. The HAT domain is represented in cyan, the bromodomain in magenta.

maps. I will try to further improve the model by measuring seleno-methionine (SeMet) derivatives and calculating anomalous difference map to locate missing methionines.

At the moment, a partial structure of the p300 'core' is available (about 80% of completeness, 120 aa missing over 580 aa per monomer) **(Figure 46)**. The region of the density map belonging to the two BRDs and to the two HAT domains present in the asymmetric unit could be assigned. Extra-density map for the PHD domain could be seen between the two separate domains. However, this connecting density was not defined enough to allow the modeling of the latter domain. Inspection of this preliminary, partial p300 'core' model showed that the BRD substrate recognition site is located in close proximity to the active site of the HAT domain, suggesting a role for this module in p300 substrate targeting **(Figure 47)**. The choice between the two symmetry mate bromodomains was done considering orientation and distance constraints. As the crystals of the p300 'core' were only obtained after complete deacetylation of the enzyme, the model presented here could represent the inhibited conformation of the enzyme. There is no structural information about the PHD finger region so far. Positioning of the PHD domain within the structure would shed light on the role of this domain in p300 substrate targeting and in the acetyltransferase activity regulation.

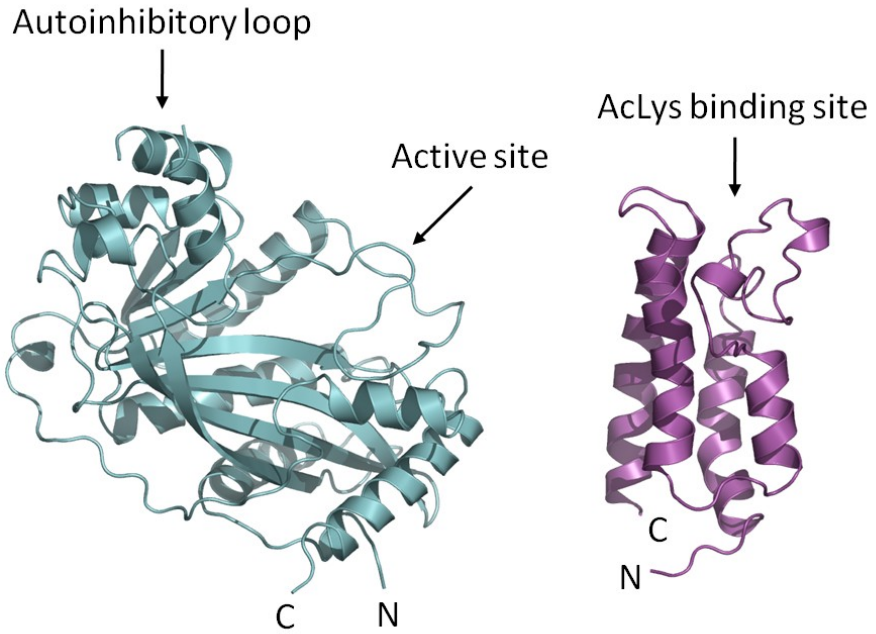


Figure 47 **Partial model of the p300 'core' monomer.** The bromodomain substrate recognition site and the HAT active site are indicated by arrows. The putative position of the autoinhibitory loop is also indicated. The loop is not present in this model because the crystals were obtained with the deacetylated p300 'core' Δ1520-1580/Y1467F construct. The HAT domain is represented in cyan, the bromodomain in magenta. N, N-terminus; C, C-terminus.

3.3.2. Low resolution model of the p300 'core' by SAXS

For SAXS measurements, several protein constructs were made. The p300/CBP 'core' Δ1520-1580/Y1467F was purified in the acetylated form. The deacetylated form of the protein was then obtained as described in Materials and Methods (Chapter 2.9.1). The HAT domain alone and the unit bromo-PHD (BP) were also purified and analyzed. SAXS measurement were performed in order to obtain a model of the p300/CBP 'core' in solution from the superimposition of three SAXS envelopes (p300/CBP 'core', BP and HAT) and of the crystal structure of the p300 HAT domain. Moreover, I wanted to investigate conformational changes upon autoacetylation of the p300 'core'.

3.3.2.1. Advantages of using liquid-chromatography-coupled SAXS

Initial SAXS data were collected on ID14-3 at the ESRF in Grenoble. The experimental apparatus consisted of a sample changer connected to the SAXS sample cell made from a quartz capillary. Scattered photons were then detected by a PILATUS pixel detector placed behind the flight tube. The sample was continuously flowing during measurements to prevent radiation damage **(Figure 48)**.

For HAT and p300 'core' samples some aggregation was observed as judged by the

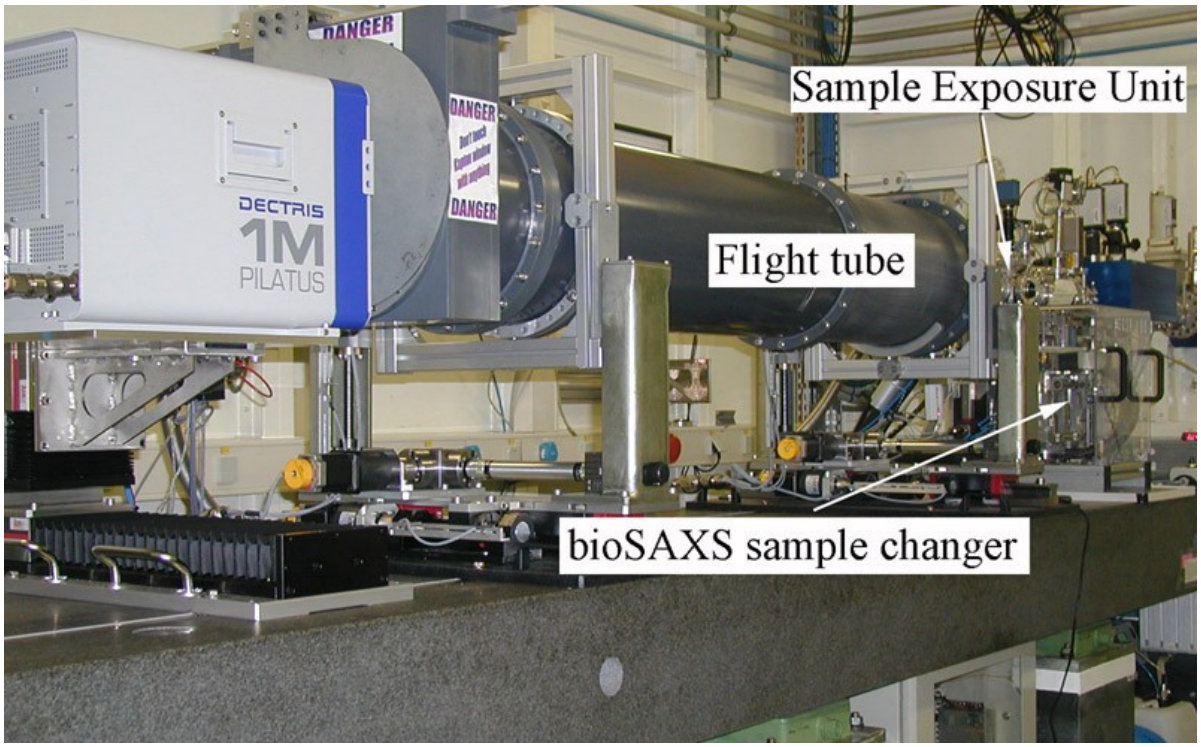


Figure 48 **Experimental apparatus of the ID14-3 beamline at the ESRF.** The bioSAXS beamline apparatus is indicated in white boxes. Details are described in the text.

slightly upward curvature of the Guinier plot at very low angles. Usually, in the case of reversible oligomerization this problem can be solved by collecting data at different concentrations and extrapolating to a concentration of zero. However, for the p300 'core' and HAT constructs it was not possible to obtain interpretable results, due to the presence of aggregation. Such behaviour is often seen when protein samples have the tendency to aggregate. Even though these samples behaved well during sample analysis by dynamic light scattering and size exclusion chromatography I could not prevent this problem. Every effort in order to remove aggregates, such as centrifugation or further purification directly before performing the SAXS experiment was not sufficient to overcome this problem.

One of the techniques developed to allow the analysis of SAXS data from samples that show aggregation behaviour consists in collecting SAXS data using an in-line FPLC size exclusion chromatography step to remove protein aggregates right before the measurement **(Figure 49)** [137]. This setup allows removal of data frames containing aggregated sample post data collection. This provides an overall more homogeneous dataset as compared to experimental setups that measure scattering from all particles in solution. It is important to note that with this setup the sample is also continuously flowed during the measurement, reducing problems of radiation damage.

An experimental setup that includes an in-line FPLC is available at the SWING beamline at the SOLEIL synchrotron in Paris **(Figure 50)**. The p300 constructs were measured at SWING in order to overcome the problem of aggregation encountered in the preliminary studies carried out at the ESRF. The protein samples prone to aggregate were

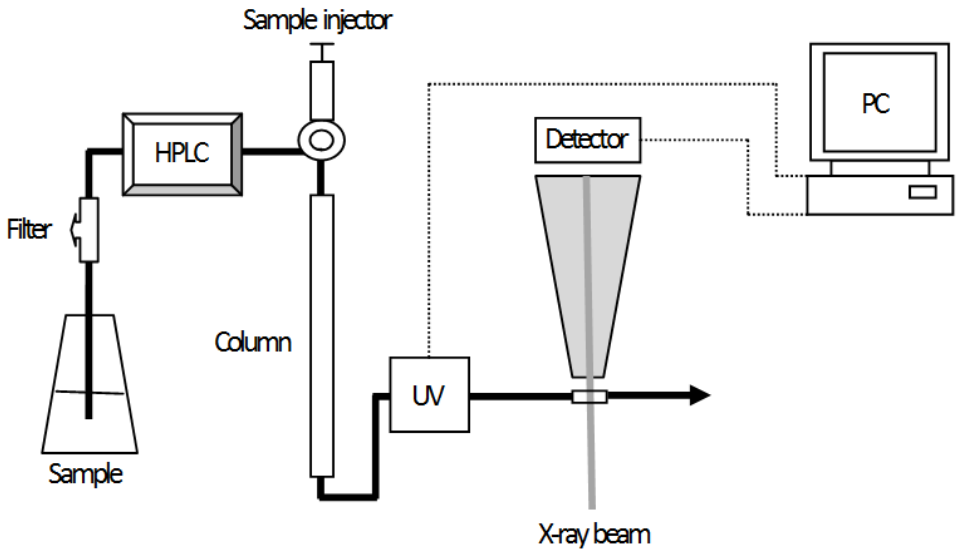


Figure 49 **Schematic view of a liquid-chromatography-coupled SAXS measurement.** The sample is passed through a gel filtration column and monitored for UV absorption immediately before the sample chamber. Scattered photons are then recorded on the detector.

eluted from the gel filtration column in a single peak. Inspection of individual data frames across the peak allowed for the discrimination of the different oligomeric species within the elution peak. The data corresponding to the monomeric species alone could then be retained and used for further calculations and p300 modelling.

It is important to add that the error of concentration measurement of the macromolecules in solution is one of the biggest issues in the calculation of protein size using SAXS (on average 20% of error estimate). At SWING, this problem is avoided by using an UV detector prior the SAXS sample chamber. This allows precise measurement of the protein concentration in the elution peak based on the absorbance at 280 nm.

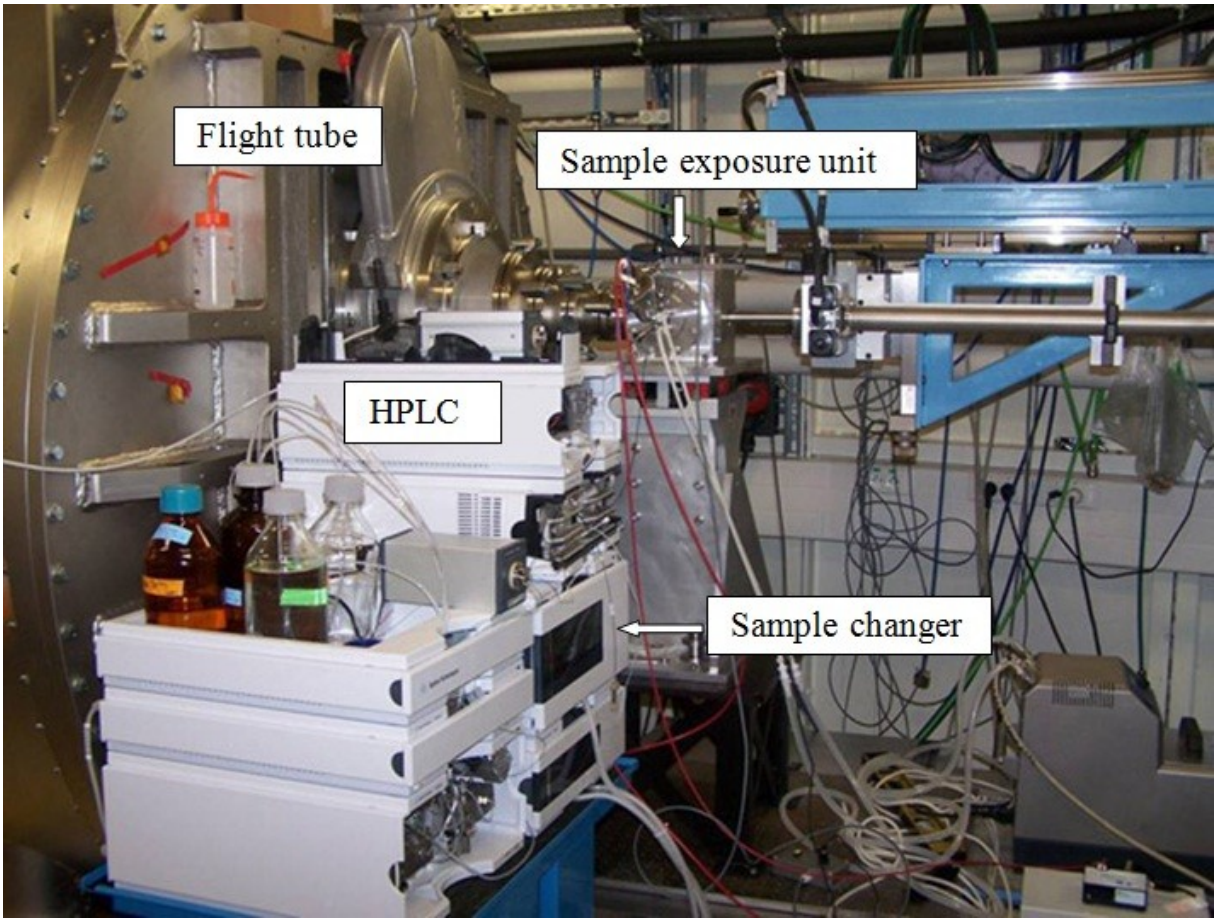


Figure 50 **Experimental setup on the beamline SWING at the SOLEIL synchrotron.** A gel filtration chromatography setup (HPLC) is coupled with the SAXS apparatus.

3.3.2.2. Determination of the p300 'core' structure by SAXS

Data collected at the SWING beamline allowed characterization of the p300 constructs. The BP module eluted like a single monomeric species from the gel filtration column and the entire peak was considered for data processing. In the case of the HAT and the p300 'core' constructs, the areas of the peak corresponding to protein aggregates were excluded from the study, as described by Mathew *et al.* [137]. The aggregation status of the sample could be determined by observing the behaviour of the R_g (radius of gyration), which remained constant across the part of the peak containing the monodisperse sample. Signals from higher molecular weight impurities or aggregates could be detected due to their characteristic low I(0) and high R_g values and removed from further analysis.

It was possible to obtain good fit to the experimental data upon processing with GNOM **(figure 51A)**. From the plot of the interatomic distance distribution function it was possible to derive information about dimensions of the particles in solution. As shown in **Figure 51B**, the BP module and the HAT domain constructs are globular, whereas the p300 'core' constructs are of elongated shape. This is in agreement with the crystal structure of the HAT domain, which is globular, and with the partial X-ray structure of the p300 'core', which has an elongated shape.

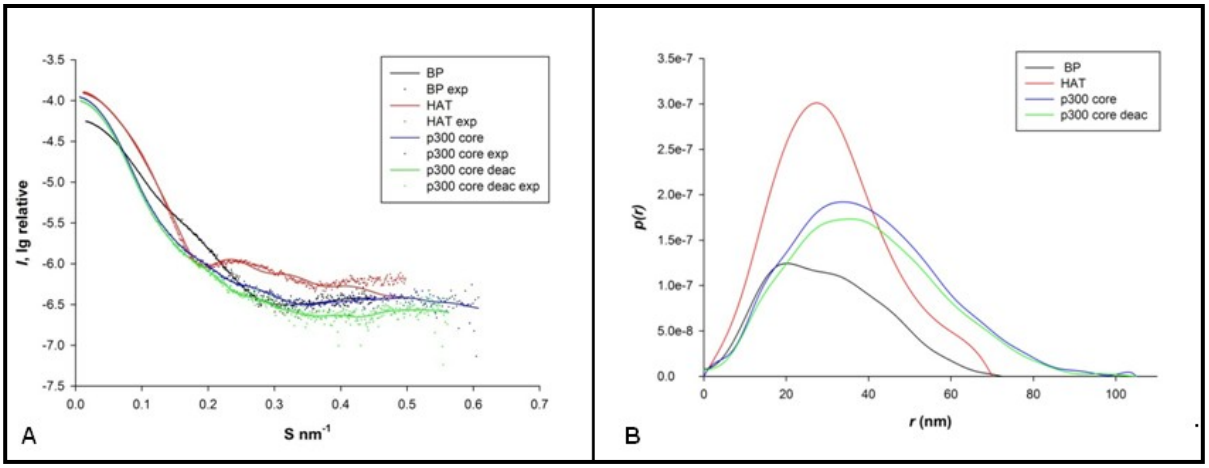


Figure 51 **Analysis of p300 constructs by SAXS using GNOM.** (A) The experimental SAXS profile. Log intensities are plotted as function of momentum transfer. Dots: experimental; lines: calculated. (B) Normalized interatomic distance distribution functions computed from the experimental scattering patterns. The colour code is the same in (A) and (B): black, BP module; red, HAT; blue, p300 'core'; green, deacetylated p300 'core'.

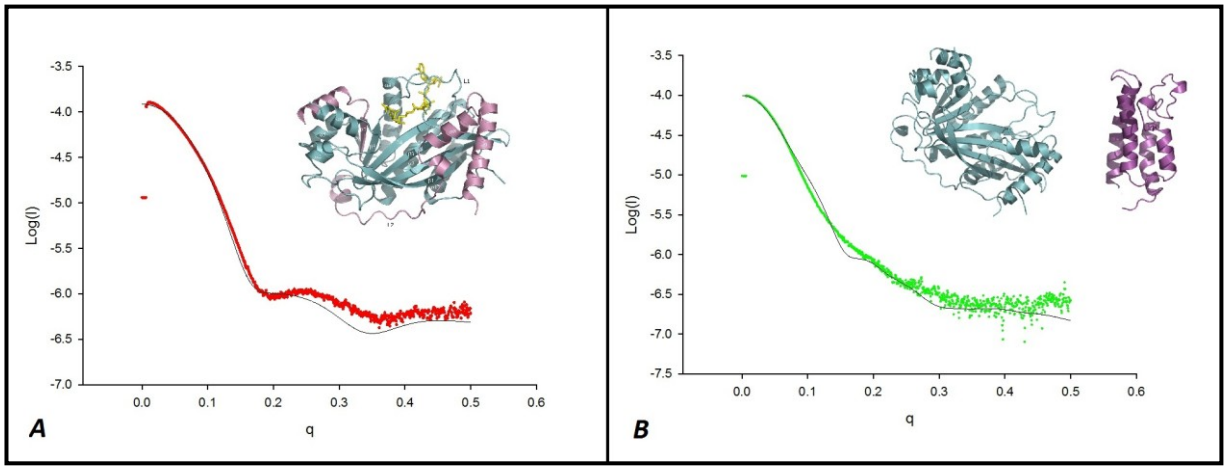


Figure 52 **SAXS curves computed by CRYSOL.** Comparison of I(q) calculated by CRYSOL from the atomic structure of the HAT (A) and the partial p300 'core' model (B). The Log of the intensities is plotted as a function of the scattering vector (q). Experimental data: red (HAT) or green (p300 'core') dots; calculated curves: black line.

Parameters obtained for the p300 constructs are reported in **Table 7**. The R_h (hydrodynamic radii) measured by DLS agreed very well with the R_g (radii of gyration) obtained by SAXS. Moreover, the molecular mass calculation by SAXS matched very well with the theoretical one.

To validate the SAXS results, theoretical intensities for high-resolution models were calculated using CRYSOL and compared to the experimental scattering profile. The HAT domain structure (PDB, ID 3BIY) fitted well to the experimentally measured data ($\chi 2 = 7.6$) **(Figure 52A)**. The partial model of the p300 'core' also yielded reasonably good fits ($\chi 2 = 2.3$) **(Figure 52B)**.

Table 7. **Overall parameters of p300 from DLS, SLS and SAXS.** *Rh*, *Rg*, *Dmax* are, respectively, the hydrodynamic radius, experimental radius of gyration and maximum size. *MMexp* , *MMsls* and *MMth* and molar masses determined by SAXS and SLS and calculated from the primary sequences, respectively. *Ch* and *Csls* are the concentration of samples used for DLS and SLS, respectively.

	DLS	SLS	SAXS					
Sample	R_h, nm (C_h, mg·ml$_{-1}$)	MM_{SLS} kDa (C_{SLS}, mg·ml$_{-1}$)	R_g nm	D_{max} nm	MM_{exp} kDa	MM_{th} kDa	Guinier I(0)	Conc [mg/ml]
BP	2.61 (1)	n/d	2.31	7.25	27.9	27.4	6.36 e-05	1.33 (peak)
HAT	n/d	n/d	2.38	7	37.5	37.9	1.69 e-04	1.72 (peak)
Core	3.35 (1)	63.8 (2)	3.06	10.5	67	67	1.61 e-04	1.40 (peak)
Core_deac	n/d	n/d	3.05	10.5	62.6	67	2.14 e-04	1.12 (peak)

A comparison of the D_{max} and the R_g obtained for the acetylated and deacetylated p300 'core' constructs showed that there were no major differences between the two states **(Table 7)**. *Ab-initio* modelling done using DAMMIF and DAMAVER and rigid body modeling using SASREF showed that the overall arrangement of the three domains (BRD, PHD finger and HAT) was not changing upon deacetylation **(Figure 53)**. In summary, the SAXS data shows that the crystallized HAT domain and p300 'core' maintain a similar structure in solution and that autoacetylation of the p300 'core' does not result in major conformational rearrangements.

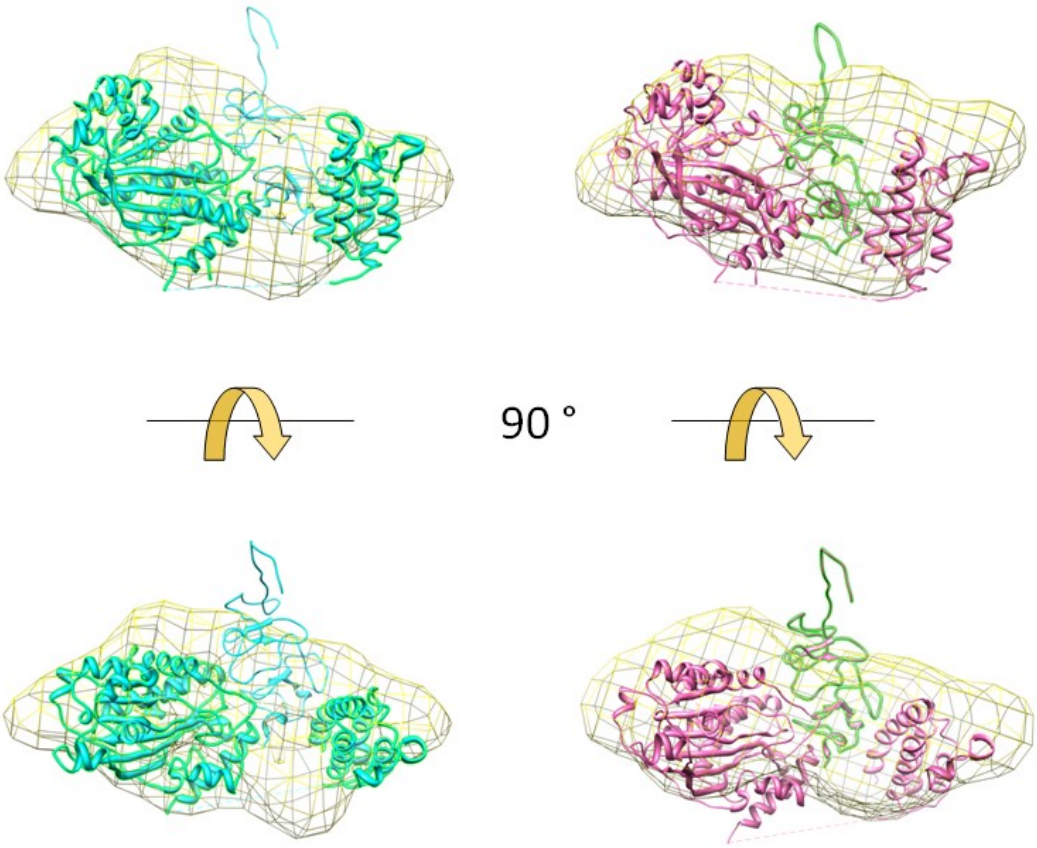


Figure 53 **Low resolution model of the p300 'core'.** Superposition of the low resolution *ab-initio* models of the p300 'core' (yellow) with the models of the p300 'core' hypoacetylated (cyan) and deacetylated (pink) obtained by rigid body refinement. For the rigid body refinement, the partial crystal structure of the p300 'core' and an homology model of the PHD domain were used. For convenience, the PHD domain is represented in a different colour in both structures. In the model, from left to right are: the HAT domain, the PHD finger, the bromodomain.

3.4. CHARACTERIZATION OF THE BP MODULE

3.4.1. ITC binding studies with histone peptides

According to our hypothesis, the BP module might serve as a substrate targeting module for p300. It has been previously reported that the BP module is involved in targeting p300 to substrates [88]. However, BP substrate specificity has not been investigated in detail. For this purpose, several post-translationally modified histone H3 and H4 peptides were tested for binding to the BP module using ITC.

Mono- and di- acetylated histone H3 and H4 peptides were tested in order to determine the binding specificity of the p300 BRD. Binding parameters are reported in **Table 8** and **9**.

Table 8. Thermodynamic analysis of the interaction between BP and 20-mer histone H3 peptides. Experimental errors were calculated to be about 2% in the number of binding sites (*n*), max 20% in the binding enthalpy (ΔH) and max 20% in the dissociation constant (K_d). Sequences: ARTKQTQRKSTGGKAPRKQL (1-20), TGGKAPRKQLATKASRKSAP (11-30).

Peptide H3		*n*	ΔH (kcal/mol^{-1})	K_d (μM)
(1-20)	K14K18 diac	1.03	-3649	104
(1-20)	K14 ac	1.13	-1.33 e^{-4}	1761
(1-20)	K18 ac	1.01	-1.35 e^{-4}	1988
(11-30)	K18K23 diac	No binding	---	---
(11-30)	K23 ac	No binding	---	---
(1-20)	K9K14 diac	1.06	-2965	578
(1-20)	K9 ac	No binding	---	---
(1-20)	K4 me1	No binding	---	---
(1-20)	K4 me3	No binding	---	---
(1-20)	S10 pho	No binding	---	---
(1-20)	T11 pho	No binding	---	---
(1-20)	No PTMs	No binding	---	---

In case of modifications on histone H3, the BP module could discriminate between mono- and di- acetylated substrates. In particular, the binding was preferential for di-acetylated substrates carrying the consensus sequence K(Ac)-XXX-K(Ac). The three amino acids in the centre were preferentially Gly or Ala. A good example is the H3K18K23 di-acetylated peptide, which was not bound by the BP module because of the presence of four

aminoacids between the two modified lysines. Moreover, this peptide was the only one carrying a modification on the residue H3K23 which is not a p300 preferential substrate *in vivo* and *in vitro*. Therefore, the ITC experiments also suggested that the BP module could only bind to substrates that were modified by p300.

In agreement with the data collected on the H3 peptides, BP was shown to preferentially bind di-acetylated H4 peptides which contained a spacing of three small amino acids between the two modified lysines. The H4 peptides tested that are bound by the BP module contain modifications that have been reported to be catalyzed by p300 both *in vivo* and *in vitro*. In general, the BP module showed in general higher affinity for H4 modified peptides as compared to the H3 peptides. In summary, data collected on histone H3 and H4 peptides showed that the BP module binds to acetylated substrates through its BRD and that di-acetylated substrates are preferred. The consensus binding sequence is: K(Ac)-XXX-K(Ac). Moreover, the BP module can bind modifications which are catalyzed by p300.

I also attempted to identify histone substrates that bind to the PHD domain **(Table 8)**. To this end, some diversely modified H3 peptides were tested. First, H3 peptides mono- and tri- methylated on K4 were tested but no binding was detected. I then tested if the BP module binds to H3 peptides phosphorylated on Ser10 and Thr11. However, no binding was observed. I also attempted to use histone peptide microarrays (Alta Bioscience) in order to detect novel BP substrates. However, none of these trials resulted in unambiguous results. So far, the role of the p300 PHD domain in p300 chromatin recognition remains unclear.

Table 9 Thermodynamic analysis of the interaction between BP and 21-mer histone H4 peptides. Experimental errors were calculated to be about 2% in the number of binding sites (*n*), max 20% in the binding enthalpy (ΔH) and max 20% in the dissociation constant (K_d). Sequence: GKGGKGLGKGGAKRHRKVLRD (4-24)

Peptide H4	*n*	ΔH (kcal/mol^{-1})	K_d (μM)
K5 ac	No binding	---	---
K8 ac	1.04	-898	58
K5K8 diac	1.03	-6312	828
K8K12 diac	1.01	-4809	90
K12 ac	1.05	-6121	71
K16 ac	No binding	---	---
K12K16 diac	0.99	-3647	25
No PTMs	No binding	---	---

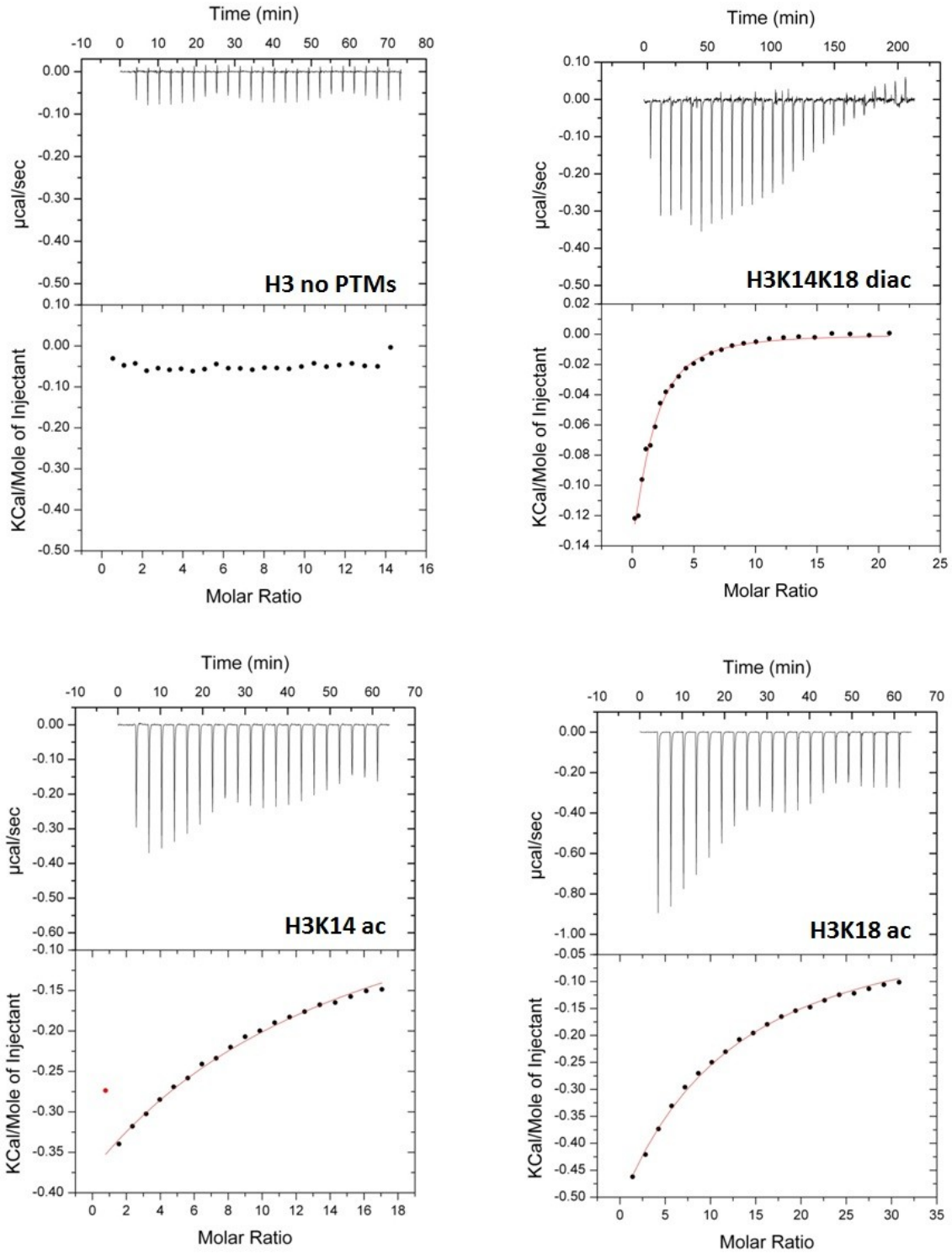


Figure 54 **Calorimetric titration of H3 peptides with the BP module.** Black dots represent the experimental data, continuous red lines correspond to the best fitting to a model considering one set of binding sites. The names of the 20-mer peptides tested are indicated.

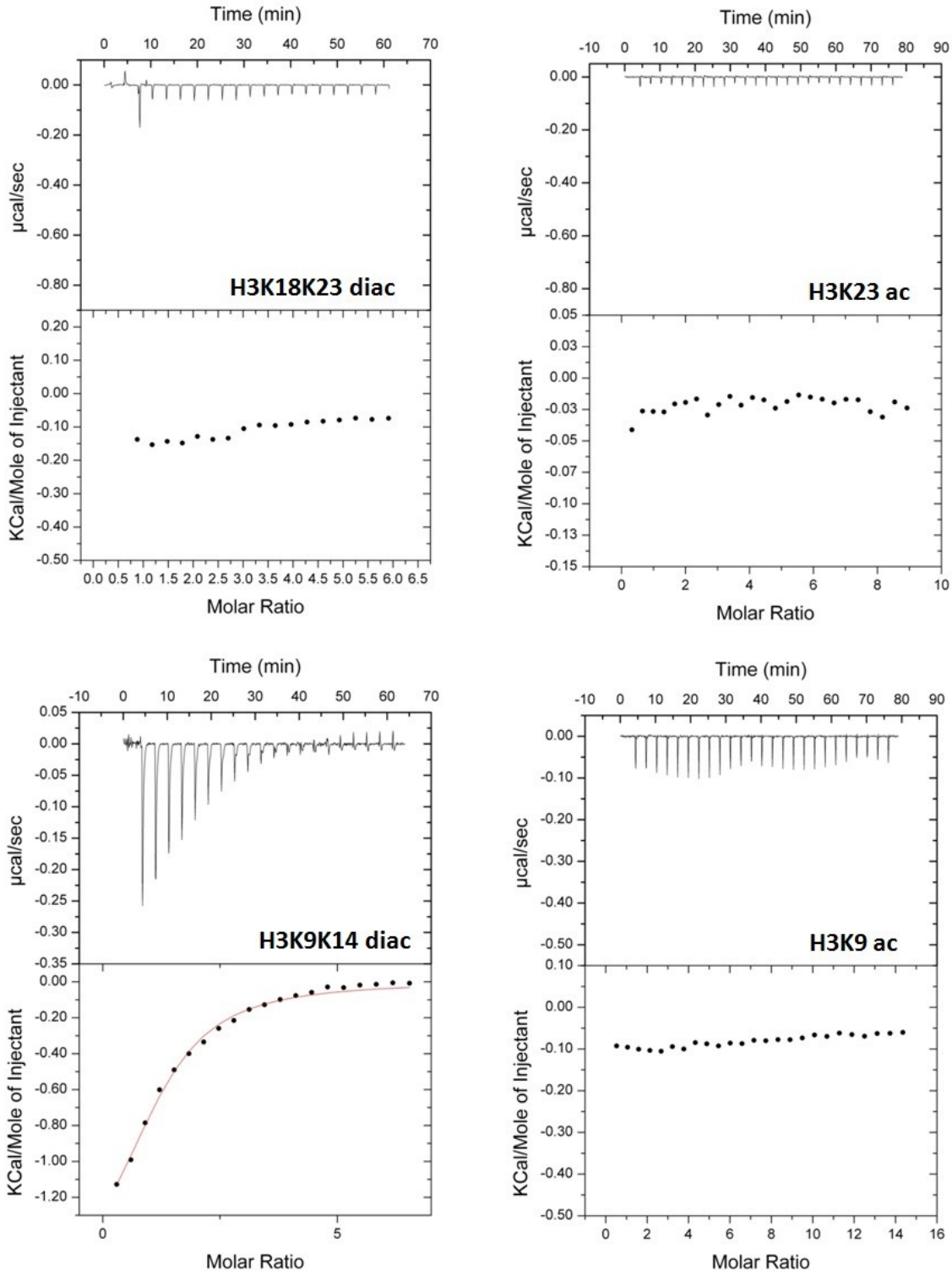


Figure 55 **Calorimetric titration of H3 peptides with the BP module.** Black dots represent the experimental data, continuous red lines correspond to the best fitting to a model considering one set of binding sites. The names of the 20-mer peptides tested are indicated.

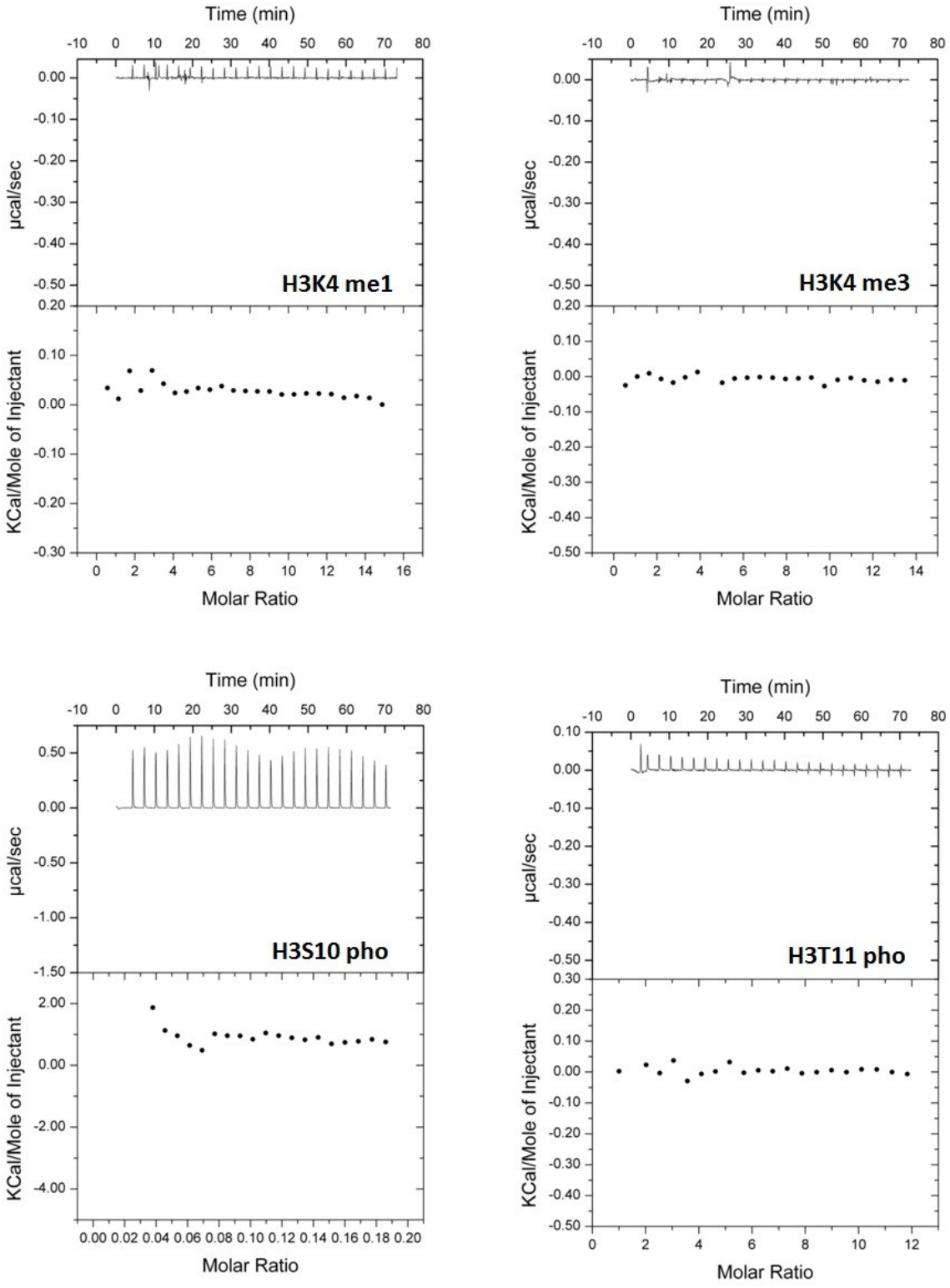


Figure 56 **Calorimetric titration of H3 peptides with the BP module.** Black dots represent the experimental data. The names of the 20-mer peptides tested are indicated.

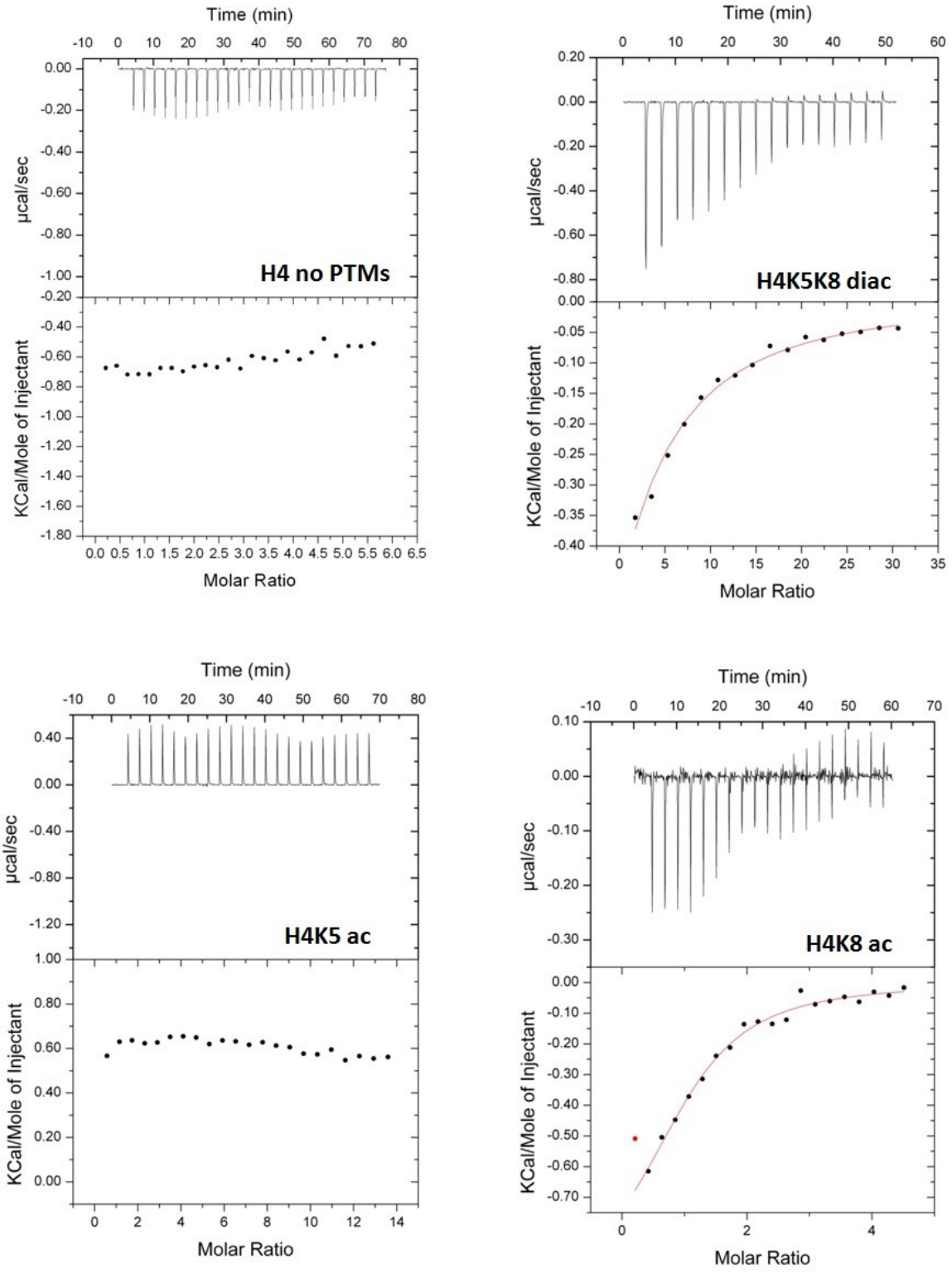


Figure 57 **Calorimetric titration of H4 peptides with the BP module.** Black dots represent the experimental data, continuous red lines correspond to the best fitting to a model considering one set of binding sites. The names of the 21-mer peptides tested are indicated.

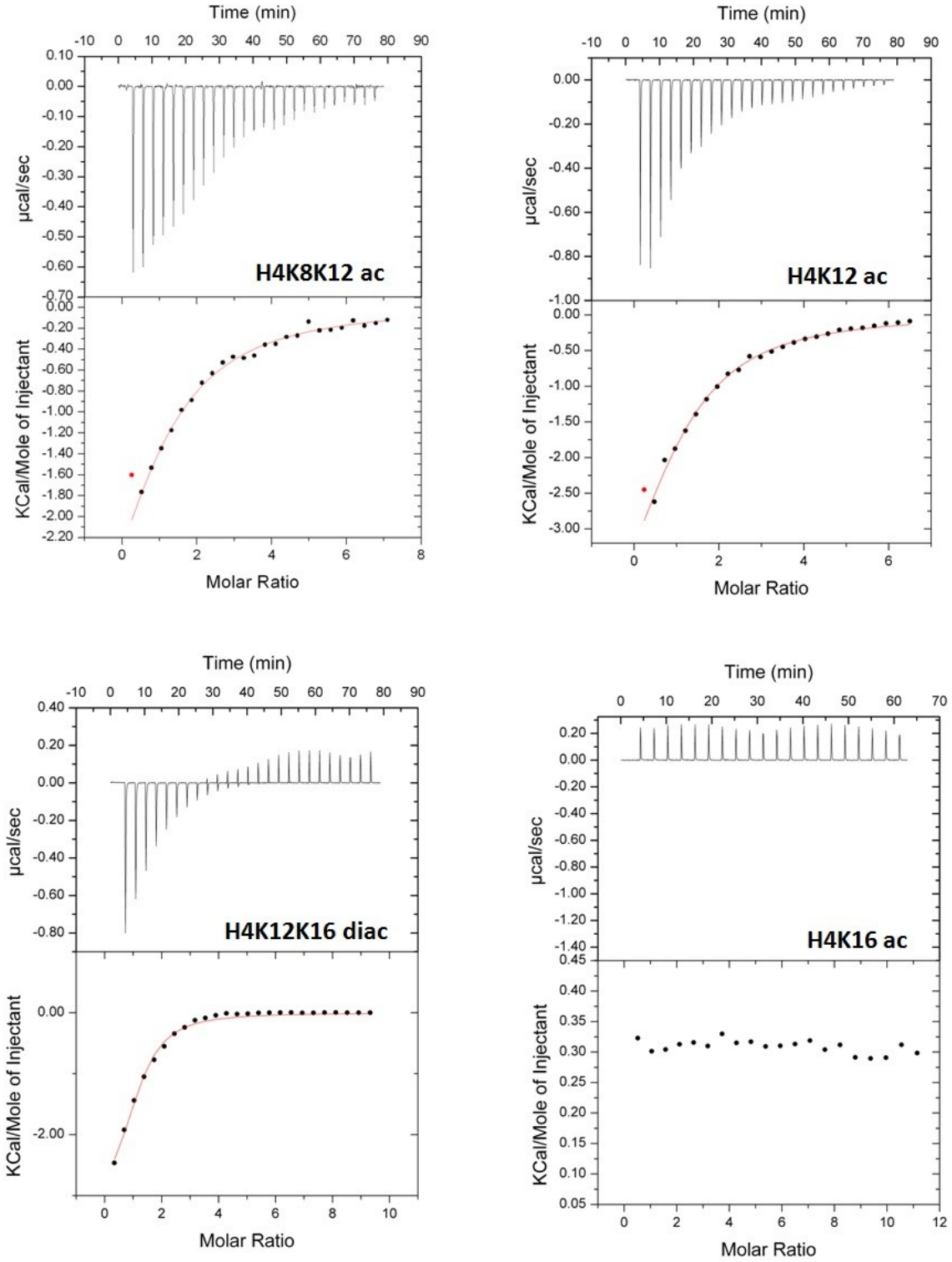


Figure 58 **Calorimetric titration of H4 peptides with the BP module.** Black dots represent the experimental data, continuous red lines correspond to the best fitting to a model considering one set of binding sites. The names of the 21-mer peptides tested are indicated.

3.4.2. BP crystallization

Crystals of the p300 BP module were obtained using purified protein cleaved off the tag. Crystals grew as clusters of needles at 4°C in ammonium sulphate (1.5-1.7 M) and Bis/Tris or Hepes buffers (pH 6.0-7.0) **(Figure 59A)**. A number of different techniques were used in order to improve crystal quality, as described in Chapter 2.9.3. However, it was not possible to produce single crystals of good quality suitable for structural studies. Several needles and some small single crystals (20-30 μm) were tested for diffraction on the microfocus beamline ID23-2 (ESRF) and were shown to be protein crystals. However, it was not possible to obtain good diffraction data.

The BP module of the p300 homolog CBP was also expressed and purified in order to get structural information. Crystals grew under a single condition (0.1 M Tris/HCl pH 8.5, 0.2 M NaCl and 25% PEG3350, at 4°C) and were reproducible. Crystal size and shape was of better quality as compared to p300 BP crystals **(Figure 59B)**. CBP BP crystals were tested on ID14-4 (ESRF) and were shown to be protein crystals. However, as for the p300 BP crystals, it was not possible to obtain good quality diffraction data.

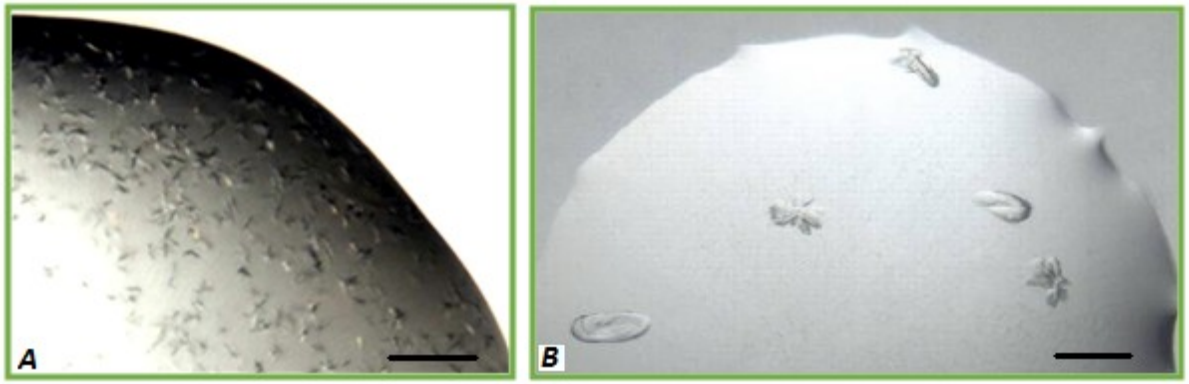


Figure 59 **Crystals of the BP module.** (A) Crystals of p300 BP. (B) Crystals of CBP BP. Black lines correspond to 0.1 mm.

4. DISCUSSION

RÉSUMÉ EN FRANÇAIS

Dans le chapitre suivant les résultats obtenus sont discutés par rapport au modèle proposé. En particulier, il a été possible de démontrer que le module BP n'est pas seulement impliqué dans la régulation de p300, mais aussi dans son ciblage aux substrats. Cette régulation est due à l'interaction entre le module BP et le domaine HAT. La structure partielle du 'noyau' de p300 montre que le bromodomaine n'interagit pas directement avec le domaine HAT. Puisque la fonction du domaine PHD n'a pas encore été clarifiée, il est possible que ce domaine soit responsable du lien avec le domaine HAT et de la régulation de l'activité enzymatique de p300. Il est important de noter que la structure actuelle du 'noyau' de p300 montre l'enzyme dans son état hypo-acétylé et inactif.

Afin de déterminer la spécificité du module BP pour des peptides d'histones modifiés post-traductionnellement, des études ont donné des informations importantes sur la séquence consensus reconnue par le bromodomaine. En particulier, le bromodomaine de p300 est le deuxième connu jusqu'ici, après Brdt-BD1, capable de reconnaître simultanément deux acétylations sur le même substrat. Contrairement à Brdt-BD1, qui reconnaît les lysines acétylées et séparées par seulement deux aminoacides, le bromodomaine de p300 nécessite la présence de trois résidus d'acide aminé entre le deux lysines acétylées.

Jusqu'à présent, il y avait plusieurs exemples où des combinaisons particulières de module sont impliqués dans l'autorégulation des protéines liés à la chromatine. Dans ce chapitre, deux exemples particuliers sont décrits. Dans le premier, l'un des modules impliqué dans la régulation a une activité différente par rapport à la normale. Dans le second, les deux modules impliqués dans la régulation ont une activité normale de reconnaissance des MPTs mais l'utilisent de manière combinatoire.

Le premier exemple est MLL1 (*Mixed Lineage Leukaemia 1*), une méthyltransférase qui contient la cassette PHD3-bromo. L'isomérisation d'un résidu de proline dans la région de liaison entre le domaine PHD3 et le bromodomaine sert de "commutateur" entre le statut actif et inactif de MLL1. En outre, le bromodomaine présent dans la protéine n'est pas capable de reconnaître des lysines acétylées. Sa fonction est de stabiliser la liaison du domaine PHD3 avec ses substrats lorsque la proline est dans sa conformation *trans*, ou d'empêcher l'accès à certains substrats lorsqu'elle est en position *cis*.

Le deuxième exemple est rsc4, une sous-unité du complexe RSC, qui contient un double bromodomaine. Dans ce cas, le deux bromodomaines conservent leur capacité à reconnaître les lysines acétylées. Le mécanisme de régulation exige le lien intramoléculaire de l'un de deux domaines avec un résidu auto-acétylé de la protéine. Ce lien entraîne l'établissement d'un état répressif.

Il est possible que le module BP de p300 puisse fonctionner de manière semblable à l'un des modules décrits dans les exemples précédents. Par la suite, des hypothèses sur la façon dont la régulation de p300 pourrait être mise en œuvre par le module BP seront présentées dans le deux cas.

4. DISCUSSION

4.1. ROLE OF THE BROMO-PHD MODULE IN HAT REGULATION

The p300 acetyltransferase is a very important transcriptional coactivator, whose misregulation has severe impact on cellular viability and proliferation. This renders the study of p300 regulation very important for medical research. Regulation is thought to occur in the central catalytic 'core' of the protein. This 'core' contains the histone acetyltransferase domain (or HAT) and the bromo-PHD (BP) module.

In order to investigate the role of the BP module on HAT regulation, a HAT construct containing a deletion of the autoinhibitory loop was used for activity assays on histone peptides (Chapter 3.1.2 and Chapter 3.2.2). The statistics calculated from the activity assays in absence and in presence of BP were subjected to significant standard deviation. However, the kinetic parameters obtained in absence of BP were similar to the ones previously obtained with another HAT domain construct deleted of the loop [6, 92]. A comparison of the parameters obtained for the mutant as well as for the *wt* protein is presented in **Table 10**. In general, the double reciprocal plot suggests a mechanism of mixed inhibition of the HAT domain by the BP module (Chapter 3.2.2).

Mixed inhibitors usually regulate reactions by binding to the enzyme (or to the enzyme-substrate complex) affecting the enzyme-substrate complex formation. It is important to note that the inhibitor usually binds to a site which is distinct from the enzyme active site. In case of p300, I could demonstrate that the BP module binds tightly to the HAT domain and that this interaction is stable (Chapter 3.2.1). This interaction may occur through a binding surface which is different from the p300 active site. Indeed, the autoinhibitory loop present in the HAT domain has been proposed to inhibit the acetyltransferase activity upon binding to the pocket P2 within the active site. This pocket is highly electronegatively charged and is thought to accommodate the lysine-rich hypoacetylated autoinhibitory loop [92]. The binding of the BP module to the HAT domain may represent an additional level of autoregulation, which would help maintaining the enzyme in the inactive state.

I introduced point mutations in the BP module, affecting key residues of the bromodomain (N1131) and of the PHD domain (K1253). Using these mutants, I could show

that the BP module is not binding to the HAT domain through the recognition sites usually used by BRDs and PHD fingers in the interaction with PTMs (Chapter 3.2.1). More data should be collected in order to characterize the BP-HAT interaction surface. In particular, from the crystal structure of the p300 'core' it will be possible to design point mutations affecting the BP-HAT binding surface and to test these mutants for *in vitro* and *in vivo* activity (see Chapter 5.3).

It is possible that the PHD domain, whose role has not yet been determined, may be a major player involved in the BP-HAT interaction. The preliminary model of the p300 'core' suggests that there is no direct interaction between the BRD and the HAT domain. Therefore, it is possible that PHD domain (that is currently missing in the model) mediates contacts between the BP module and the HAT domain.

An interesting hypothesis is that post-translational modifications, such as autoacetylation, of the p300 'core' regulate BP-HAT interaction and serve as a switch between active and inactive state of the enzyme. Limited proteolysis shows subtle differences between the differently acetylated p300 'core' preparations that are suggestive of conformational differences (Chapter 3.1.1). Alternatively, as the protease (trypsin) used in these experiments cleaves C-terminal to Arginine and Lysine, autoacetylation on Lysine residues could cause changes in the cleavage pattern. It would be interesting to test this model further by doing similar experiments on the deacetylated p300 'core'.

There are additional observations that suggest that autoacetylation does not directly regulate BP-HAT interaction. In pull-down assays, the BP module interacts with the HAT irrespective of degree of acetylation of the HAT domain (chapter 3.2.1). SAXS studies on the deacetylated and partially autoacetylated p300 'core' yielded very similar data showing that there are no major conformational rearrangements. In summary, data presented here suggests that the BP module constitutively binds to the HAT domain and that autoacetylation does not regulate this interaction.

Table 10. Comparison of kinetic constants for HAT catalyzed acetyl transfer.

Construct	Peptide	K_m (μM)	K_{cat} (s^{-1})	Reference
HAT Δloop	H3	4.39 ± 3.08	0.92 ± 0.45	This work
HAT Δloop	H4	24 ± 7	0.53 ± 0.05	Thompson *et al.* (2004) [6]
HAT *wt*	H4	162 ± 77	0.26 ± 0.05	Thompson *et al.* (2004) [6]
HAT *wt*	H4	160 ± 10	4.1 ± 0.1	Liu *et al.* (2008) [92]

4.2. STRUCTURAL CHARACTERIZATION OF THE p300 'CORE'

The best way to characterize the interaction between the BP module and the HAT domain and the conformational changes occurring during the transition of p300 between its inactive and active states would be to have structures of the p300 'core' in the different states and in presence of the autoinhibitory loop. The knowledge of an accurate p300 structure can also be very important in rational drug design, since structure-based functional studies can help in the development of more effective and specific therapeutic agents and therapies.

The intrinsic toxicity of p300 constructs when overexpressed in host cells is one of the biggest problems for all research groups who are trying to approach the molecular characterization of p300. Upon overexpression in host cells, the p300 acetyltransferase is toxic because it indiscriminately acetylate host proteins. This results in low yields and production of heterogeneously modified p300. Moreover, p300 is composed of several structured domains which are connected by unstructured regions making it difficult to obtain stable, well-behaved protein constructs. Several approaches have been used in order to overcome the difficulties in handling recombinant p300, such as the co-expression with deacetylases and the production of semi-synthetic constructs [6]. Here, I developed a novel approach for p300 production that resulted in successful p300 structural characterization.

First, a construct containing the conserved 'core' region of the protein (aa 1048-1666) was designed. A catalytic point mutation (Y1467F) and a deletion of the autoinhibitory loop were also introduced (Chapter 3.1.1). Together, these modifications allowed good protein overexpression in *Hi5* insect cells but did not lead to crystallizable p300 'core'. Therefore, low-resolution structural studies were performed using SAXS.

It was shown by SAXS that the p300 'core' has an elongated shape and that there are no major conformational changes within the p300 'core' during the transition between different acetylation states (Chapter 3.3.2). These changes may be too small to be detected by SAXS and may not require large domain rearrangements. It is then very unlikely that big conformational changes occur upon acetylation. The movement of a small loop within the PHD domain, together with the already known displacement of the autoinhibitory loop upon autoacetylation, might be enough for p300 activity regulation.

The crucial approach that finally led to p300 crystallization was the complete deacetylation of the p300 'core' Δloop/Y1467F by SIRT2 to obtain homogeneous protein preparations. Crystals of p300 'core' were obtained using deacetylated protein prepared as described in Chapter 2.9.1. As proposed in Chapter 1.6 and Chapter 4.4, the crystallized form of the p300 'core' might represent an inactive conformation of the enzyme. As the

autoinhibitory loop was deleted in the construct used for structural analysis, it was not possible to see if it was bound to the p300 active site when hypoacetylated. However, space proximity suggested that this interaction could indeed be possible. I attempted to co-crystallize the p300 'core' with a autoinhibitory loop peptide (residues 1545-1560). Crystals and diffraction data were obtained but failed to show additional density in the resulting map.

From the partial reconstruction of the p300 'core' structure it was possible to note that the BRD does not make direct contacts with the HAT domain, suggesting that the main surface of interaction might be with the PHD finger. Unfortunately, due to the poor quality of the electron density map for this region of the protein, it was not possible to build a model for the p300 PHD finger and to determine the surface of interaction with the HAT. In general, all PHD fingers are inherently flexible domains and this renders very difficult their characterization using X-ray crystallography. Indeed, the method which is usually used for their structural characterization is NMR.

It is possible that the p300 PHD domain has a function which is other then being a regulatory domain that mediates the binding to the HAT domain. For instance, it could be involved with the BRD in p300 substrate targeting (multivalent engagement of different PTMs by linked binding modules) or it could simply serve as a structural domain to maintain the HAT domain and the BRD in the correct orientations for activity.

As mentioned before, the partial X-ray model of the p300 'core' shows that the p300 BRD is not binding directly to the HAT domain. However, it is positioned very close to the HAT domain. The BRD substrate recognition site (constituted by the loops ZA and BC which are opposite to the N- and C- termini of the domain) is oriented in the same direction as the HAT active site. This suggests that the BRD may serve as substrate targeting module for p300, instead of being involved in the BP-HAT interaction, like discussed above.

Having the atomic structure of p300 would help to understand some of the point mutations associated with p300 misregulation and disease appearance. For instance, mutations affecting the p300 HAT domain were previously studied by structure-based point mutagenesis to determine the impact on protein autoregulation [92]. Using this approach, Liu *et al.* were able to explain the role of some HAT point-mutations in the appearance of severe disorders, such as particular carcinomas. Similarly, point mutations occurring within the BP module can be the cause of the development of several disorders, such as the Rubinstein-Taybi Syndrome, breast and colorectal carcinomas [2, 5]. Structure based analysis will allow for the interpretation of these mutations. For instance, it is possible that the presence of a single point mutation can destabilize the BP-HAT interaction, preventing the autoregulatory action of the BP module.

4.3. ROLE OF THE BROMO-PHD MODULE IN p300 TARGETING

Isothermal Titration Calorimetry (ITC) studies were performed in order to measure the interaction between the BP module and modified histone peptides. Using this biophysical technique it was possible to determine the binding parameters (*n*, K_a, ΔH and ΔS) for BP-peptide interaction (see Chapter 2.7 and Chapter 3.4). The ITC studies on modified histone peptides allowed understanding of some of the basis of BP substrate recognition. In particular, the role of the BRD in recognizing AcLys was confirmed, but it was not possible to determine the binding specificity of the PHD domain.

Interestingly, the p300 BRD could bind to monoacetyl-lysine weakly and this interaction was significantly reinforced when a second acetyl-lysine was present at 3-aa distance. As an example, BP bound the diacetylated peptide H3K14K18 (K(Ac)XXX-K(Ac)) but not the diacetylated H3K18K23 (K(Ac)-XXXX-K(Ac)). It must be mentioned that the BP could also bind to peptides which were acetylated on lysine residues separated by 2 or 4 amino acids; however the affinity was weaker and comparable to that of mono-acetylated peptides. The 3-aa linker between the two modified residues contains preferentially small amino acids. Presumably, the presence of bulky amino acids in the linker leads to steric hindrance. For example, in the case of the diacetylated peptide H4K8K12 the presence of the hydrophobic leucine might be the reason why the interaction with BP was slightly weaker than in the case of the monoacetylated H4K8 and H4K12 peptides. Moreover, the generally higher affinity of the BP module for the histone H4 peptides over the histone H3 ones could be explained considering their aminoacid composition: the H4 tail is highly rich in amino acids with small side-chains, such as glycine and alanine. Some other techniques were used to confirm the data obtained by ITC, such as NMR and histone peptide microarrays. However, the BP module resulted not to be suitable for these techniques.

ITC binding experiments performed with the BP N1131A mutant (mutation within the BRD) showed no binding for the same acetylated peptides normally recognized by the BP (data not shown). This suggested that the single BRD in p300 is able to recognize diacetylated peptides. The idea that the p300 BRD could recognize di-acetylation marks was already suggested by sequence comparison with the BRD Brdt-BD1, known to recognize diacetylated peptides [68]. Morinière *et al.* showed that in Brdt-BD1 the binding pocket defined by the loops ZA and BC is wider than that of canonical BRDs, such as the Gcn5 BRD. This allows binding of two acetyl-lysine residues in the same pocket. Structure alignment of the Brdt-BD1 and CBP/p300 BRDs revealed that the latter ones are sterically compatible with the

recognition of diacetylated peptides [68]. Moreover, binding studies carried out by Wei and colleagues also showed that the p300 BRD interacts with a di-acetylated MyoD peptide [116]. Both groups proposed that the p300 BRD binds di-acetylated peptides and that the spacing between the two acetylated lysines may be of 2-aa. The first hypothesis was confirmed by the ITC experiments reported here. However, according to my data, the p300 BRD can bind very weakly acetylated lysines separated by 2-aa and instead prefers a spacing of 3-aa between the two modified residues. In summary, the p300 BRD represents the second reported BRD capable of simultaneously binding two acetylated marks on a histone tail.

It is important to note that Wei and colleagues also demonstrated that the BRD of p300 was capable of binding specifically di-acetylated marks. The BRD of CBP instead discriminates poorly between mono- and di-acetylated MyoD peptides [116]. This finding suggested that p300 and CBP might not have completely interchangeable role in cellular processes. This was also suggested by Morinière and colleagues after structure alignment of the Brdt-BD1 and the CBP/p300 BRDs [68]. Of the four BRD residues that distinguish p300 from CBP, only one is present in the active site of the proteins: Tyr 1141 in p300 which is substituted by Phe 1177 in CBP. Superposition with Brdt-BD1 suggests that the tyrosine residue within the p300 BRD binding site might form an extra hydrogen bond with the backbone carbonyl of a second acetylated lysine of the diacetylated peptide [68]. According to this observation, the difference between the p300 and CBP BRDs in the single residue Tyr1141 (p300) *versus* Phe 1177 (CBP) could explain the different ligand specificity of the two proteins **(Figure 60)**. Co-crystal structures of the p300 and CBP BRDs with modified peptides would reveal the molecular bases of substrate recognition and specificity of the two

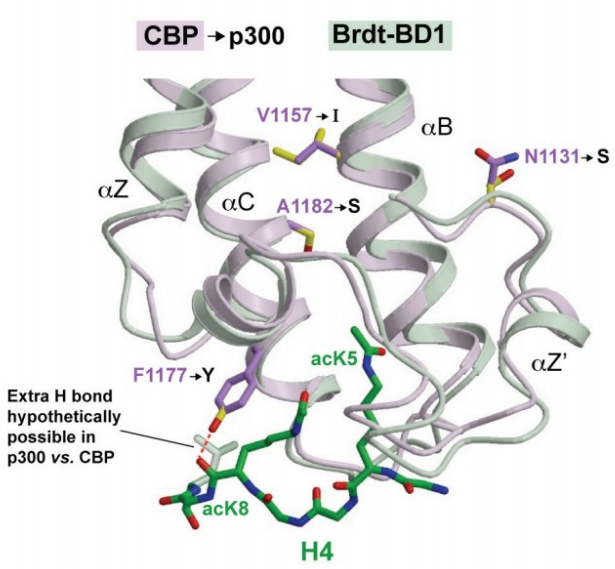


Figure 60 **Superposition of the Brdt-BD1 and the CBP/p300 bromodomains.** Superposition of Brdt-BD1 (PDB: ID 2WP2; light green) and CBP (PDB: ID 2RNY; magenta). The side chains of the four residues that distinguish p300 from CBP are shown with their carbon atoms in yellow (p300) and magenta (CBP) [68].

proteins.

An interesting finding also coming from the ITC binding studies is that the BP module is capable of recognizing substrates which are modified by p300. All modifications on histone H4 peptides that were found to bind the BRD are p300 substrates *in vivo* and *in vitro* [106]. It is known that p300 can preferentially modify H3K14 and H3K18 *in vitro* and that knock-down of p300/CBP dramatically and specifically reduce H3K18 and H3K27 acetylation *in vivo,* whereas H3K23 acetylation is not affected [106, 138]. Among the modified histone H3 peptides tested, BP could bind H3K14K18 diacetylated peptides, but it was not able to bind either monoacetylated H3K23 or diacetylated H3K18K23 peptides. As K23 is not a site preferentially subjected to p300-mediated post-translational modification, the lack of binding by the BP module to this site suggests that BP exclusively binds diacetylated p300 substrates, with high preference for the K(Ac)-XXX-K(Ac) sequence motif. This indicates that p300 histone modification and binding may be coupled events.

4.4. THE PROPOSED MODEL

The data obtained indicates that the BRD and the PHD finger module (BP) of p300 can function not only in p300 substrate targeting but also in auto-regulation of p300 enzymatic activity. In **Figure 61**, I propose a model for the functional cycle of p300 activation and substrate binding. The idea that the BRD and the PHD module may function as single unit is suggested by the fact that other PHD-BRD combinations have already been reported to function in a concerted way, such as the PHD-BRD modules of BPTF [76] and KAP1 [87].

In the proposed model, I assume that acetylation/deacetylation events are important for p300 regulation, as suggested by the studies carried out by Thompson and colleagues on the autoinhibitory loop [6]. I also assume that the p300 'core' is hypoacetylated in its inhibited/inactive state. It has been proposed that p300 inhibition is in part due to the action of the lysine-rich autoinhibitory loop. When p300 is in its inactive hypoacetylated state, this loop is though to be bound to the P2 pocket in the catalytic site of the HAT domain, impairing the access of substrates to the active site [92]. Liu *et al.* also suggested that the loop would loose affinity for the P2 pocket upon autoacetylation and be released from the active site removing inhibition of p300 [92].

I propose that binding of the BP module to the HAT domain also inhibits the enzyme. This binding would not occur in the active site of the protein (mixed inhibition) and thus would not interfere directly with substrate binding. According to the X-ray analysis of the

p300 'core', the PHD finger appears to mediate the interaction between the HAT domain and the BP module, without direct interactions between the BRD and the p300 catalytic site. The proposed model can be described as follows **(Figure 61)**.

Upon trans(autoacetylation) the autoinhibitory loop looses affinity for the binding to the catalytic site of the protein and p300 activity is stimulated. The binding between the BP module and the HAT domain does not appear to be influenced by acetylation events occurring on the HAT domain. It is possible that acetylation or other modifications within the BP module destabilize the interaction and serve as a switch to relieve the inhibitory action of the BP. Indeed, as revealed by mass spectrometry, some autoacetylation events can target residues present within the BP module. However, the conformational changes that occur during the transition between inactive and active states (detected by limited proteolysis and SAXS studies) do not seem to require major domain rearrangements. These changes probably consist in local rearrangement of flexible regions, such as the autoinhibitory loop. The interaction between the HAT domain and the BP module could also be weakened by loss of surface contacts because of, for instance, post-translational modifications of key residues within the BP module or rearrangements of loops within the PHD domain.

Once activated, p300 can start modifying key residues on target substrates. As the BP binds substrates which have been already modified by p300 (as demonstrated by ITC), it likely targets other p300 molecules to regions of chromatin already p300 modified, leading to an increase in local enzyme concentration and activity. The result would be the establishment

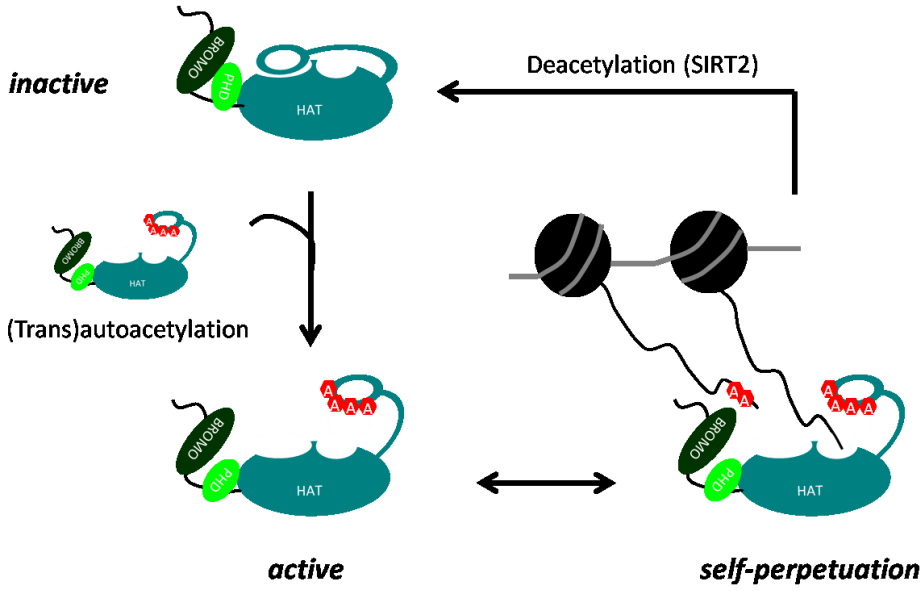


Figure 61 **Proposed model for p300/CBP 'core' regulation.** Cycles of acetylation/deacetylation events are important for p300/CBP activity, as well as bromo- and PHD- domains. In the figure are shown the bromodomain (bottle green), the PHD domain (apple green) and the HAT domain (navy green). Acetylated lysines are indicated by red hexagons. Nucleosomes: black circles; histone tail: black line; DNA: gray line.

of a positive feedback loop of p300-dependent acetylation events which would maintain the particular targeted region of the chromatin in a stable hyperacetylated state. When the need of p300 acetylation is decreased, the protein may finally undergo deacetylation and go back to the inactive state, terminating the cycle.

The implication of the BP module in p300 regulation can be very important to understand some of the disorders associated with p300 misregulation. Particular types of cancers, such as breast and gastric carcinomas, and diseases, such as the Rubinstein-Taybi Syndrome, are indeed associated with the loss or the mutation of the BRD or the PHD domain. Since the BP module also appears to be implicated in maintaining the inactive-inhibited state of p300, any alteration that mutates this module may lead to p300 misregulation. In this context, the structure of the BP module and of the p300 'core' could be very informative for the development of new therapies and the design of more effective and specific drugs.

4.5. AUTOACETYLATION AND REGULATION OF HATS

As mentioned before, Thompson *et al.* have shown that acetylation in the autoinhibitory loop of p300 stimulates protein activity in a similar way as phosphorylation of the inhibitory loop in protein kinases [6]. Autoacetylation of HATs may be a common mechanism for regulation of acetyltransferases activity. In addition to p300, several HATs have been shown to self-acetylate *in vitro* [90]. For these enzymes, acetylation usually stimulates HAT activity, even though the mechanism of activation and regulation seems to diverge.

For instance, a close homolog of p300, the Rtt109 (regulation of Ty1 transposition protein 109) acetyltransferase was recently reported to be regulated by autoacetylation [139]. The structure of Rtt109 was solved using X-ray crystallography [139-141]. Overall, the structures of the p300 and the Rtt109 HAT domain are very similar. Both structures are characterized by the presence of an extended, 7-stranded β-sheet that is surrounded by several α-helices. Moreover, the unusually long L1 and L2 loops are conserved in both proteins and the loop L1 is essential for AcCoA cofactor binding [139-141] **(Figure 62A-B)**.

Despite the structural similarities between Rtt109 and p300, several properties of the two enzymes suggest that they diverge significantly at the functional level and in the mechanism of autoregulation. The surface charge of Rtt109 proximal to the lysine binding site is apolar, similar to that in the GNAT and MYST family, which is in contrast with the

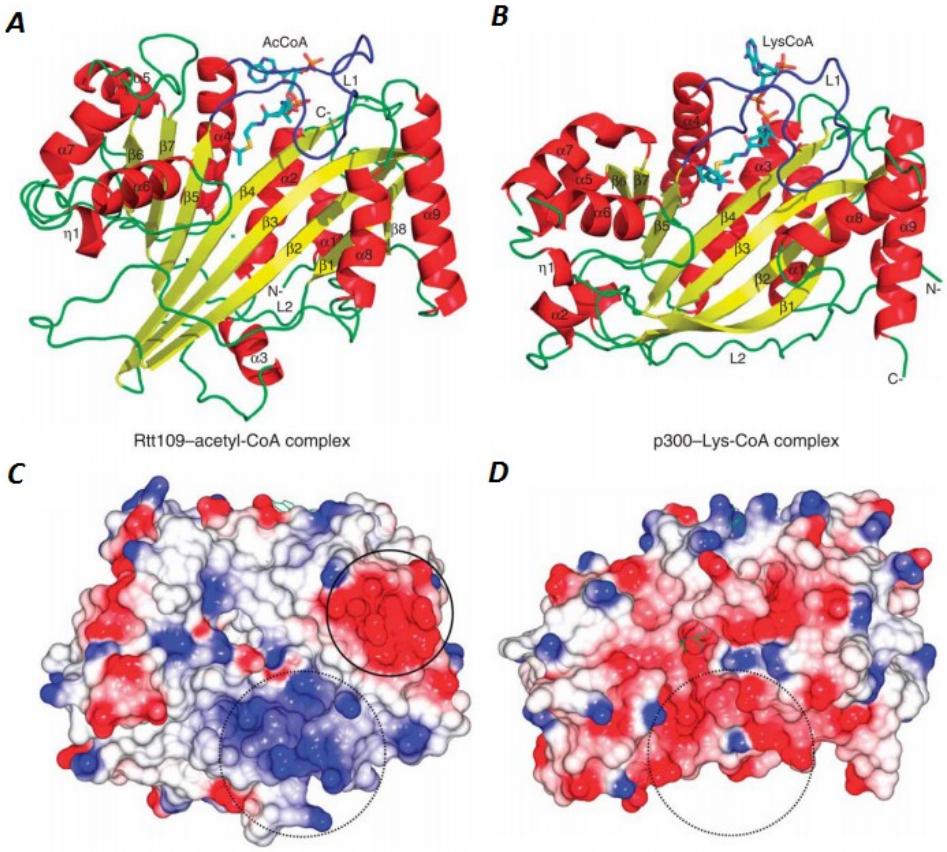


Figure 62 **Structure of the Rtt109 acetyltransferase and comparison with p300.** (A-B) Crystal structure of Rtt109 bound to AcCoA (A) and p300 bound to LysCoA (B) are shown. Secondary structural elements are numbered and represented in yellow (β-sheets) and red (α-helices). The AcCoA cofactor and LysCoA inhibitor are shown in sticks representation. (C-D) Electrostatic surface representation of the Rtt109 HAT domain (C) and of the p300 HAT domain (D). Open circles indicate the region proximal to the lysine binding pocket. Closed circle represent an acidic patch conserved in Rtt109 homologs [140].

electronegative surface of p300 [140] **(Figure 62B-C)**. Moreover, Rtt109 is very selective for H3K59 and H3K9, unlike p300 which has a broad substrate specificity. Rtt109 is also catalyzing acetyl-transfer through a mechansims which differs from that of other HATs, since neither the conserved glutamate residue present in the GNAT and MYST family nor the two p300 catalytic residues (Trp 1436 and Tyr1467) are present [121].

As mentioned above, Rtt109 was shown to autoacetylate [139, 142]. Autoacetylation is occurring intramolecularly and only on one lysine residue located 8 Å away from the substrate binding site of the protein (Lys 290) [139, 142]. This suggests that the mechanism of activation by autoacetylation is different from that of p300, which requires rapid and intermolecular acetylation of 7-10 lysines within the autoinhibitory loop for activation [7]. In the case of Rtt109, it was reported that acetylation of K290 induces a structural change within the protein active site, which results in improved binding of AcCoA and enhances the catalytic transfer of the acetyl group onto the lysine substrate [142].

Rtt109 represents the first detailed example of a HAT enzyme regulated by a single, intramolecular autoacetylation event. Like Rtt109 and p300, other HATs seem to undergo self-acetylation. For instance, the presence of a structurally relevant lysine acetylation within the HAT domain is common in many mammalian members of the MYST family. In the case of MOF (male absent on the first), a member of the MYST family of HATs, acetylation of a lysine residue (K274) is important to maintain structural integrity of the HAT domain and to allow the correct positioning of conserved residues involved in substrate binding [143].

Overall, it is becoming clear that autoacetylation and subsequent enzyme activation may be a common regulatory mechanism for many acetyltransferases.

4.6. REGULATORY MECHANISMS BY OTHER EFFECTOR MODULES

It is possible that the BP module can take part in regulation of the p300 'core' in several different ways. On one hand, the PHD finger could stabilize the BRD and enhance its binding to the diacetylated peptides, stabilizing a positive feedback loop. On the other hand, the BP module could be post-translationally modified by another enzyme or by p300 autoacetylation and this modification could serve as a switch between its inhibitory and targeting functions. Moreover, the PHD domain of p300 seems to be essential for enzymatic activity, since the replacement of this domain with homologues ones results in inactive p300 constructs [10]. The atypically large loop 1 present within this domain may be one of the players implicated in p300 regulation.

To date several examples of chromatin-related proteins have been reported to be regulated by effector modules. The p300 bromo-PHD module is likely to be another example of this type and it might have something in common with the other ones. There are mostly two types of linked binding modules: those that carry out new functions not possessed by the two individual domains and those where the two domains retain their normal function as PTMs recognition modules but act in a concerted way.

An example of the first type is the MLL1's (Mixed Lineage Leukaemia 1) PHD3-bromo cassette [144]. MLL1 is a methyltransferase which can both activate and repress genes during haematopoiesis. This enzyme can activate transcription by binding through its PHD3 domain to H3K4me3 and can repress it upon interaction with the RNA recognition motif (RRM) of the cyclophilin CyP33. The mechanism which directs MLL's binary function is dependent on the PHD3-bromo cassette. A linker region connecting the BRD and the PHD finger contains a proline residue which can undergo isomerisation The BRD function is to

enhance the binding of PHD3 to substrates when the linker proline is in the *trans*-configuration and to inhibit the access of the RRM (but not of the H3K4me3) to the PHD3 binding site when the proline is in the *cis*-configuration **(Figure 63)**. Proline isomerisation thus represents the switch between active and inactive conformation of MLL1. It is important to note that in the case of the MLL1's PHD3-bromo cassette, the BRD carries out a function which is atypical for this kind of effector module. Indeed, this BRD is not able to bind acetylated lysines. [144].

In the case of the p300 bromo-PHD module, the SAXS analysis indicates that major conformational changes like those occurring in the case of MLL1 are very unlikely to happen upon autoacetylation. However, it would be possible to think at the PHD domain of p300 as at the BRD of MLL1. The p300 PHD domain could stabilize the BRD for its binding to acetylated substrates when the enzyme is active and then switch to its inhibitory function to maintain p300 in the inactive form. It is important to note that MLL1 is a methyltransferase which is able to modify H3K4 and to bind to the H3K4me3 through its PHD3 domain. The BRD enhances the binding of the PHD3 domain to modified chromatin. This can provide a stable association of MLL1 with its target and create a positive feedback loop that reinforces transcriptional activation. In the case of p300, the PHD domain could reinforce the association of the BRD with acetylated chromatin and mediate the establishment of a positive feedback loop. Data are required to explore this hypothesis and to investigate if the p300 PHD domain is more than a simple PTM recognition module for substrate targeting.

There are also examples of linked binding modules which can exert their typical

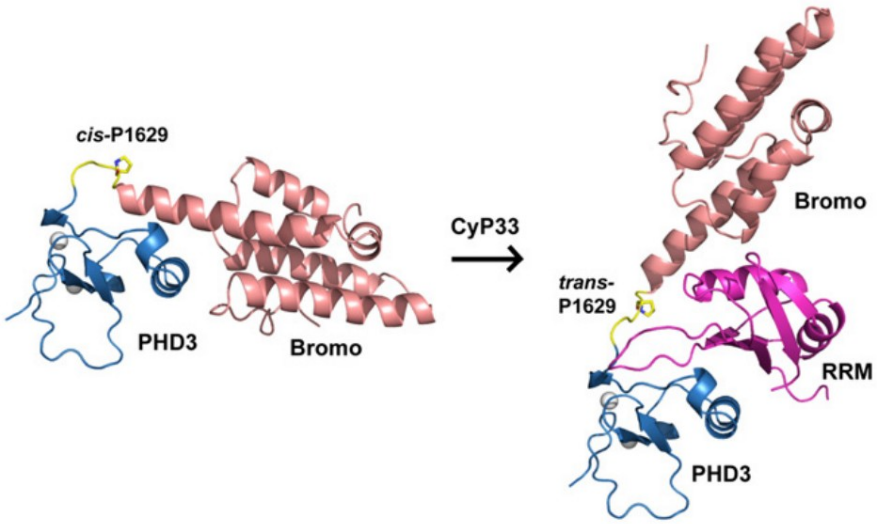


Figure 63 **The MLL1's PHD3-bromo cassette.** Structure of the MLL1's PHD3-bromo cassette with the linker Proline P1629 in the *cis*-configuration (left) and a model of the complex with the RNA recognition motif (RRM) with the linker Proline P1629 in the *trans*-configuration (right) [144].

function simultaneously and as well have an impact on protein autoregulation. A good example is represented by Src-kinases, where the cross-talk between SH2 and SH3 domains mediates substrate targeting of the active enzyme. In this case, these domains can also bind to the enzyme in absence of a substrate and maintain a stable, self-inhibited state [8]. Autoregulatory mechanisms where a protein adopts a repressed conformation in response to intramolecular binding of PTMs are commonly found in protein containing modules involved in the recognition of PTMs, also in the case of lysine acetylation. For example, the tandem BRD of rsc4 (a subunit of the RSC complex) contains two BRDs, BD1 and BD2. Acetylation of rsc4 can be recognized by the BD1 BRD and this binding prevents the second BRD from interacting with its substrates and promote transcription [86].

Similar mechanisms could as well be postulated for the p300 acetyltransferase. In particular, we can consider the fact that acetylation/deacetylation events occurring within the autoinhibitory loop in p300 were already shown to play a regulatory role similar to the one of phosphorylation/dephosphorylation for kinases [6]. The bromo-PHD module could bind internally the HAT domain stabilizing the autoinhibited conformation and be unlatched upon autoacetylation, occurring presumably within the BP module itself.

In conclusion, the p300 bromo-PHD module plays a pivotal role in protein regulation and substrate targeting. The new data from this work shed light on the mechanism of this regulation, which is now more defined. However, more experiments will need to be carried out in order to dissect in details all the steps required for BP inhibition/activation of p300 activity. Medical research will benefit largely from any progress performed in this direction.

5. CONCLUSION AND FUTURE PERSPECTIVES

RÉSUMÉ EN FRANÇAIS

Il est évident que le module BP joue un rôle clé dans la régulation et le ciblage des substrats de p300. Les mécanismes par lesquels le module BP pourrait influencer l'activité de p300 sont de mieux en mieux définies grâce aux études menées pendant cette thèse. Cependant, des études complémentaires devraient être faites afin de déterminer dans le détail les différentes étapes nécessaires à l'inhibition et l'activation de p300 par son module BP.

Le fait que ce module BP est impliqué dans la régulation de p300 conduirait, à l'avenir, à l'utiliser comme une nouvelle cible pharmacologique. En effet jusqu'à présent, la recherche de médicaments spécifiques dirigés contre les acétyltransférases a été particulièrement difficile. Le plus gros problème réside dans le fait que le degré de conservation du domaine catalytique des acétyltransférases reste très élevé et aussi que l'activité des certaines acétyltransférases est essentielle a la survie de l'organisme. En conséquence, les médicaments développés jusqu'ici sont caractérisées par une faible spécificité et une faible efficacité. La possibilité d'utiliser le module BP comme cible de p300 peut avoir certains avantages puisque le bromodomaine et le PHD domaine de p300 possèdent des caractéristiques particulières qui ne se retrouvent pas dans d'autres protéines de la même famille.

La structure de p300, une fois mieux définie, représentera une information clé pour le développement de nouvelles thérapies plus efficaces. En outre, il sera possible d'interpréter certaines mutations ponctuelles liées à la survenue de maladie graves, telles que le syndrome Rubinstein-Taybi.

A l'avenir, des efforts seront faits afin de mieux caractériser l'interaction entre le module BP et le domaine HAT. En outre, des études seront effectuées afin de rechercher le rôle du domaine PHD et de mieux déterminer la spécificité de substrat du module BP. L'impact des événements d'auto-acétylation sur la régulation de p300 sera encore plus approfondit. Enfin, compte tenu des preuves de l'implication du domaine CH3 dans la régulation de p300, d'autres études seront menées en tenant compte de cette importante donnée.

5. CONCLUSION AND FUTURE PERSPECTIVES

5.1. CONCLUSION

The p300 acetyltransferase is implicated in many biological processes and its regulation is often very important to maintain cell viability. However, the details about p300 regulation mechanisms are not yet understood. In this work, I studied the central catalytic 'core' of p300 which contains the catalytic HAT domain and the bromo-PHD module. Althoug previous studies on p300 activity and regulation have been performed [6], they focused only on the HAT domain. Using the p300 'core' construct, I showed that the p300/CBP acetyltransferase activity is controlled by an autoregulatory mechanism mediated by the bromo-PHD module. I found that the bromo-PHD module acts as a mixed inhibitor for p300 activity upon binding to the HAT domain. I also determined that this binding is acetylation independent and that it may be mediated by the PHD domain.

Moreover, the crystal structure of the p300 'core' suggests that the bromodomain functions as a substrate-targeting module for p300. ITC binding studies confirmed the ability of the p300 bromodomain to bind acetylated residues. Very importantly, the p300 bromodomain was shown to specifically recognize di-acetylated substrates and to bind substrates which are already modified by p300. This result suggests that the p300 BP module could serve to stabilize the interaction of p300 with hyperacetylated chromatin and to promote the establishment of positive feedback loops.

Taken together, these data demonstrated for the first time that, similar to other chromatin related proteins, the action of linked binding modules is pivotal for p300activity regulation. This finding represents an important tool for future medical research and for the development of more specific drug design.

5.2. IMPACT OF THE STUDY

Structure-based drug design of small inhibitor molecules specific for the HAT domain of acetyltransferases is now one of the most followed approach in the design of anti-cancer drugs [145]. However, due to the high conservation of the central catalytic core in all acetyltransferases and to the fact that some HAT knock-downs are lethal, many of the newly generated HAT inhibitors tend to be characterized by low potency, low permeability or low specificity. The identification of the autoregulatory BP module in p300/CBP can be helpful for the development of more specific inhibitors. For example, novel compounds could be developed that destabilize the interaction between the BP module and the HAT domain. Moreover, both the bromodomain and the PHD domain of p300/CBP bear characteristics within their putative binding domains which are peculiar for this acetyltransferase. The bromodomain has an insertion of two aminoacids within the ZA loop, whereas the PHD domain presents a long insertion within the loop 1. For this reason I decided to perform structural studies on both the p300 'core' and the BP module alone. The partial crystal structure of the p300 'core' reported here is very promising and, once optimized, it will represent an important tool for any structure-based drug design for p300/CBP.

It is important to note that the bromodomain of p300/CBP is very unique. Indeed, it represents the second reported bromodomain (after Brdt-BD1 [68]) that is known to recognize double acetylated marks on histone substrates. This characteristic could be reflective of a role of the bromodomain in targeting stably hyperacetylated chromatin, which is in line with other biochemical studies reported on the BP module [88]. As suggested before, the direct coupling of "reading" and "writing" acetylation marks may stabilize the association of p300/CBP with its targets and provide a positive feedback loop that localizes and reinforces transcriptional activation.

In conclusion, the mechanism of function of the BP module starts to be better defined after the studies carried out during this thesis work. I demonstrated that the bromo-PHD module of p300/CBP plays a central role in protein regulation and substrate targeting. The involvement of p300/CBP in a wide range of biological processes indicates the importance and the difficulties of studying in details this acetyltransferase. Given the severe consequences of p300/CBP misregulation and the large amount of diseases which are found to be linked with them, medical research will largely benefit from any advancements performed in the study of this acetyltransferase.

5.3. FUTURE PERSPECTIVES

In order to better define the role of the BP module in p300 regulation, the first step will be to improve the quality of the crystal structure of the p300 'core'. To this end, efforts will be done in order to obtain good quality data on derivatized protein (SeMet) crystals. After the structure is available, point mutants will allow characterization of the interaction surface between the BP module and the HAT domain. In this context, it will be interesting to test point mutations and insertions in the BP module associated with the appearance of some known diseases. A good approach will be to insert these modifications in a full-length p300 construct and detect the impact on transcriptional activation and targeting to chromatin *in vivo*. It will be possible to detect the impact on transcriptional activation *in vivo* using a gene reporter assay similar to the one described in the Supplementary of the article by Reynoird *et al.* **(Appendix)**. This gene reporter assay takes advantage of the fact that p300/CBP is pivotal in mediating p53-dependent cellular response by acetylating p53. The idea is to detect the transcriptional activation of p21, a p53 target gene, in presence of the p300/CBP mutants. In addition, it will be possible to study the impact of BP mutations on targeting to chromatin by transfection of the GFP- or HA- tagged enzyme mutants in Cos7 cells and detection of co-localization with modified chromatin by immunofluorescence, using an anti-histone H4 or H3 antibody, as described in Reynoird *et al.* **(Appendix)**.

A second important aspect that will be covered in future studies is the determination of the role of the PHD module in p300 regulation and substrate targeting. The crystal structure of the p300 'core' will be instrumental to elucidate the putative function of this 'core' region. For instance, the position of the long loop1 of the PHD domain within the 'core' will be determined and it will be possible to evaluate if it is involved in the binding to the HAT domain or in the inhibition of the HAT activity. Moreover, structural analysis of the PHD domain will elucidate if its putative methyl-lysine binding site is accessible for binding to substrates in the hypoacetylated p300 'core'. In case it is accessible, more extensive substrate binding studies will necessarily be done to define the binding specificity of the p300 PHD domain. Since this binding specificity *in vivo* might be very different in comparison with the one detected *in vitro*, the ideal will be to determine substrate binding specificity by chromatin immunoprecipitation (ChIP)-based bioinformatics analysis, using techniques such as ChIP-on-chip or ChIP sequencing. This approach will allow detecting putative combinatorial readout of post-translational modifications by the BP module. The same analysis can also be applied to constructs bearing mutations in the bromodomain or in the PHD domain binding sites, in order to define the individual specificities. In this way it will be also possible to

determine the specificity of the bromodomain *in vivo*. However, a subsequent study of the binding kinetics using biophysical techniques will be needed to finally confirm the interactions. Moreover, co-crystallization trials of the BP module with modified substrates will be done. Obtaining these structures will shed light on the atomic basis of substrate recognition by the BP module and this will be very informative for subsequent structure-based drug design. In this context, it will be important to improve the experimental set up of the activity assays described here in order to set up a well-established method for the study of p300 inhibitor molecules, as well as for the *in vitro* determination of the impact of the mutations in the BP module on p300 activity.

In order to specify in details the conformational changes involved in the transition between active and inactive state of the protein, more attempts to crystallize the p300 'core' construct containing the autoinhibitory loop will be performed. Moreover, there is increasing evidence that the CH3 domain of p300/CBP may also play an important role in protein regulation. For instance, Reynoird *et al.* **(Appendix)** have demonstrated that a protein called NUT (Nuclear protein in Testis) can hyperactivate p300/CBP by stably binding to its CH3 domain. Therefore, it is likely that additional biochemical and structural studies directed to decipher the mechanisms of p300 regulation will be carried out including this domain.

Future experiments will give more insights into p300/CBP regulation and this will help defining more effective therapies for some very severe disorders, such as the Rubinstein Taybi Syndrome and several carcinomas, for which a specific treatment is not available to date.

LIST OF FIGURES

LIST OF TABLES

BIBLIOGRAPHY

1. D.Nelson, M.C., *Lehninger - Principles of biochemistry (Fifth edition)*. 2008, W.H. Freeman and Company: New York (United States of America).

2. Kalkhoven, E., et al., *Loss of CBP acetyltransferase activity by PHD finger mutations in Rubinstein-Taybi syndrome.* Hum Mol Genet, 2003. **12**(4): p. 441-50.

3. Chan, H.M. and N.B. La Thangue, *p300/CBP proteins: HATs for transcriptional bridges and scaffolds.* J Cell Sci, 2001. **114**(Pt 13): p. 2363-73.

4. Goodman, R.H. and S. Smolik, *CBP/p300 in cell growth, transformation, and development.* Genes Dev, 2000. **14**(13): p. 1553-77.

5. Iyer, N.G., H. Ozdag, and C. Caldas, *p300/CBP and cancer.* Oncogene, 2004. **23**(24): p. 4225-31.

6. Thompson, P.R., et al., *Regulation of the p300 HAT domain via a novel activation loop.* Nat Struct Mol Biol, 2004. **11**(4): p. 308-15.

7. Karanam, B., et al., *Kinetic and mass spectrometric analysis of p300 histone acetyltransferase domain autoacetylation.* J Biol Chem, 2006. **281**(52): p. 40292-301.

8. Harrison, S.C., *Variation on an Src-like theme.* Cell, 2003. **112**(6): p. 737-40.

9. Zeng, L. and M.M. Zhou, *Bromodomain: an acetyl-lysine binding domain.* FEBS Lett, 2002. **513**(1): p. 124-8.

10. Bienz, M., *The PHD finger, a nuclear protein-interaction domain.* Trends Biochem Sci, 2006. **31**(1): p. 35-40.

11. Wade, P.A., D. Pruss, and A.P. Wolffe, *Histone acetylation: chromatin in action.* Trends Biochem Sci, 1997. **22**(4): p. 128-32.

12. Weake, V.M. and J.L. Workman, *Inducible gene expression: diverse regulatory mechanisms.* Nat Rev Genet, 2010. **11**(6): p. 426-37.

13. Trinchieri, G., *Type I interferon: friend or foe?* J Exp Med, 2010. **207**(10): p. 2053-63.

14. Chen, W. and W.E. Royer, Jr., *Structural insights into interferon regulatory factor activation.* Cell Signal, 2010. **22**(6): p. 883-7.

15. Nemeth, A. and G. Langst, *Chromatin higher order structure: opening up chromatin for transcription.* Brief Funct Genomic Proteomic, 2004. **2**(4): p. 334-43.

16. Venters, B.J. and B.F. Pugh, *How eukaryotic genes are transcribed.* Crit Rev Biochem Mol Biol, 2009. **44**(2-3): p. 117-41.

17. Li, B., M. Carey, and J.L. Workman, *The role of chromatin during transcription.* Cell, 2007. **128**(4): p. 707-19.

18. Groth, A., et al., *Chromatin challenges during DNA replication and repair.* Cell, 2007. **128**(4): p. 721-33.

19. Woodcock, C.L. and S. Dimitrov, *Higher-order structure of chromatin and chromosomes.* Curr Opin Genet Dev, 2001. **11**(2): p. 130-5.

20. Dorigo, B., et al., *Nucleosome arrays reveal the two-start organization of the chromatin fiber.* Science, 2004. **306**(5701): p. 1571-3.

21. Finch, J.T. and A. Klug, *Solenoidal model for superstructure in chromatin.* Proc Natl Acad Sci U S A, 1976. **73**(6): p. 1897-901.

22. Bednar, J., et al., *Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin.* Proc Natl Acad Sci U S A, 1998. **95**(24): p. 14173-8.

23. Schalch, T., et al., *X-ray structure of a tetranucleosome and its implications for the chromatin fibre.* Nature, 2005. **436**(7047): p. 138-41.

24. Robinson, P.J., et al., *EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure.* Proc Natl Acad Sci U S A, 2006. **103**(17): p. 6506-11.

25. Wu, C., A. Bassett, and A. Travers, *A variable topology for the 30-nm chromatin fibre.* EMBO Rep, 2007. **8**(12): p. 1129-34.

26. Tremethick, D.J., *Higher-order structures of chromatin: the elusive 30 nm fiber.* Cell, 2007. **128**(4): p. 651-4.

27. Marsden, M.P. and U.K. Laemmli, *Metaphase chromosome structure: evidence for a radial loop model.* Cell, 1979. **17**(4): p. 849-58.

28. Belmont, A.S. and K. Bruce, *Visualization of G1 chromosomes: a folded, twisted, supercoiled chromonema model of interphase chromatid structure.* J Cell Biol, 1994. **127**(2): p. 287-302.

29. Hirano, T., *At the heart of the chromosome: SMC proteins in action.* Nat Rev Mol Cell Biol, 2006. **7**(5): p. 311-22.

30. Goodbourn, S., ed. *Eukaryotic gene transcription*. 1996, Oxford University Press Inc.: New York.

31. Smith, M.M., *Histone structure and function.* Curr Opin Cell Biol, 1991. **3**(3): p. 429-37.

32. Arents, G. and E.N. Moudrianakis, *The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization.* Proc Natl Acad Sci U S A, 1995. **92**(24): p. 11170-4.

33. McBryant, S.J., X. Lu, and J.C. Hansen, *Multifunctionality of the linker histones: an emerging role for protein-protein interactions.* Cell Res, 2010. **20**(5): p. 519-28.

34. Thoma, F., T. Koller, and A. Klug, *Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin.* J Cell Biol, 1979. **83**(2 Pt 1): p. 403-27.

35. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. **389**(6648): p. 251-60.

36. Davey, C.A., et al., *Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution.* J Mol Biol, 2002. **319**(5): p. 1097-113.

37. Richmond, T.J. and C.A. Davey, *The structure of DNA in the nucleosome core.* Nature, 2003. **423**(6936): p. 145-50.

38. Dorigo, B., et al., *Chromatin fiber folding: requirement for the histone H4 N-terminal tail.* J Mol Biol, 2003. **327**(1): p. 85-96.

39. Hubner, M.R. and D.L. Spector, *Chromatin dynamics.* Annu Rev Biophys, 2010. **39**: p. 471-89.

40. Segal, E. and J. Widom, *What controls nucleosome positions?* Trends Genet, 2009. **25**(8): p. 335-43.

41. Arya, G., A. Maitra, and S.A. Grigoryev, *A structural perspective on the where, how, why, and what of nucleosome positioning.* J Biomol Struct Dyn, 2010. **27**(6): p. 803-20.

42. Cairns, B.R., *The logic of chromatin architecture and remodelling at promoters.* Nature, 2009. **461**(7261): p. 193-8.

43. Eberharter, A. and P.B. Becker, *ATP-dependent nucleosome remodelling: factors and functions.* J Cell Sci, 2004. **117**(Pt 17): p. 3707-11.

44. Yamada, K., et al., *Structure and mechanism of the chromatin remodelling factor ISW1a.* Nature, 2011. **472**(7344): p. 448-53.

45. Hargreaves, D.C. and G.R. Crabtree, *ATP-dependent chromatin remodeling: genetics, genomics and mechanisms.* Cell Res, 2011. **21**(3): p. 396-420.

46. Latchman, D.S., *Eukaryotic Transcription Factors (Fifth edition)*. 2008, Elsevier: Oxford (UK).

47. Sarma, K. and D. Reinberg, *Histone variants meet their match.* Nat Rev Mol Cell Biol, 2005. **6**(2): p. 139-49.

48. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.

49. Das, C., et al., *CBP/p300-mediated acetylation of histone H3 on lysine 56.* Nature, 2009. **459**(7243): p. 113-7.

50. Williams, S.K., D. Truong, and J.K. Tyler, *Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation.* Proc Natl Acad Sci U S A, 2008. **105**(26): p. 9000-5.

51. Hodawadekar, S.C. and R. Marmorstein, *Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design.* Oncogene, 2007. **26**(37): p. 5528-40.

52. Allis, C.D., et al., *New nomenclature for chromatin-modifying enzymes.* Cell, 2007. **131**(4): p. 633-6.

53. Wang, Y., et al., *Human PAD4 regulates histone arginine methylation levels via demethylimination.* Science, 2004. **306**(5694): p. 279-83.

54. Chang, B., et al., *JMJD6 is a histone arginine demethylase.* Science, 2007. **318**(5849): p. 444-7.

55. Lee, J.S., E. Smith, and A. Shilatifard, *The language of histone crosstalk.* Cell, 2010. **142**(5): p. 682-5.

56. Arif, M., B.R. Selvi, and T.K. Kundu, *Lysine acetylation: the tale of a modification from transcription regulation to metabolism.* Chembiochem, 2010. **11**(11): p. 1501-4.

57. Vaquero, A., R. Sternglanz, and D. Reinberg, *NAD+-dependent deacetylation of H4 lysine 16 by class III HDACs.* Oncogene, 2007. **26**(37): p. 5505-20.

58. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. **403**(6765): p. 41-5.

59. Shi, X., et al., *ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression.* Nature, 2006. **442**(7098): p. 96-9.

60. Schreiber, S.L. and B.E. Bernstein, *Signaling network model of chromatin.* Cell, 2002. **111**(6): p. 771-8.

61. Ferrell, J.E., Jr., *Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability.* Curr Opin Cell Biol, 2002. **14**(2): p. 140-8.

62. Goldberg, A.D., C.D. Allis, and E. Bernstein, *Epigenetics: a landscape takes shape.* Cell, 2007. **128**(4): p. 635-8.

63. Dodd, I.B., et al., *Theoretical analysis of epigenetic cell memory by nucleosome modification.* Cell, 2007. **129**(4): p. 813-22.

64. Taverna, S.D., et al., *How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers.* Nat Struct Mol Biol, 2007. **14**(11): p. 1025-40.

65. Marmorstein, R., *Protein modules that manipulate histone tails for chromatin regulation.* Nat Rev Mol Cell Biol, 2001. **2**(6): p. 422-32.

66. Sanchez, R. and M.M. Zhou, *The role of human bromodomains in chromatin biology and gene transcription.* Curr Opin Drug Discov Devel, 2009. **12**(5): p. 659-65.

67. Dhalluin, C., et al., *Structure and ligand of a histone acetyltransferase bromodomain.* Nature, 1999. **399**(6735): p. 491-6.

68. Moriniere, J., et al., *Cooperative binding of two acetylation marks on a histone tail by a single bromodomain.* Nature, 2009. **461**(7264): p. 664-8.

69. Mujtaba, S., L. Zeng, and M.M. Zhou, *Structure and acetyl-lysine recognition of the bromodomain.* Oncogene, 2007. **26**(37): p. 5521-7.

70. Zeng, L., et al., *Structural basis of site-specific histone recognition by the bromodomains of human coactivators PCAF and CBP/p300.* Structure, 2008. **16**(4): p. 643-52.

71. Yang, X.J. and E. Seto, *Lysine acetylation: codified crosstalk with other posttranslational modifications.* Mol Cell, 2008. **31**(4): p. 449-61.

72. Winston, F. and C.D. Allis, *The bromodomain: a chromatin-targeting module?* Nat Struct Biol, 1999. **6**(7): p. 601-4.

73. Aasland, R., T.J. Gibson, and A.F. Stewart, *The PHD finger: implications for chromatin-mediated transcriptional regulation.* Trends Biochem Sci, 1995. **20**(2): p. 56-9.

74. Ivanov, A.V., et al., *PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing.* Mol Cell, 2007. **28**(5): p. 823-37.

75. Musselman, C.A. and T.G. Kutateladze, *PHD fingers: epigenetic effectors and potential drug targets.* Mol Interv, 2009. **9**(6): p. 314-23.

76. Li, H., et al., *Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF.* Nature, 2006. **442**(7098): p. 91-5.

77. Matthews, A.G., et al., *RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination.* Nature, 2007. **450**(7172): p. 1106-10.

78. Lan, F., et al., *Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression.* Nature, 2007. **448**(7154): p. 718-22.

79. Lange, M., et al., *Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex.* Genes Dev, 2008. **22**(17): p. 2370-84.

80. Kustatscher, G. and A.G. Ladurner, *Modular paths to 'decoding' and 'wiping' histone lysine methylation.* Curr Opin Chem Biol, 2007. **11**(6): p. 628-35.

81. Bottomley, M.J., *Structures of protein domains that create or recognize histone modifications.* EMBO Rep, 2004. **5**(5): p. 464-9.

82. de la Cruz, X., et al., *Do protein motifs read the histone code?* Bioessays, 2005. **27**(2): p. 164-75.

83. Ruthenburg, A.J., et al., *Multivalent engagement of chromatin modifications by linked binding modules.* Nat Rev Mol Cell Biol, 2007. **8**(12): p. 983-94.

84. Wang, Z. and D.J. Patel, *Combinatorial readout of dual histone modifications by paired chromatin-associated modules.* J Biol Chem, 2011.

85. Zeng, L., et al., *Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b.* Nature, 2010. **466**(7303): p. 258-62.

86. VanDemark, A.P., et al., *Autoregulation of the rsc4 tandem bromodomain by gcn5 acetylation.* Mol Cell, 2007. **27**(5): p. 817-28.

87. Zeng, L., et al., *Structural insights into human KAP1 PHD finger-bromodomain and its role in gene silencing.* Nat Struct Mol Biol, 2008. **15**(6): p. 626-33.

88. Ragvin, A., et al., *Nucleosome binding by the bromodomain and PHD finger of the transcriptional cofactor p300.* J Mol Biol, 2004. **337**(4): p. 773-88.

89. Marmorstein, R., *Structure of histone acetyltransferases.* J Mol Biol, 2001. **311**(3): p. 433-44.

90. Sterner, D.E. and S.L. Berger, *Acetylation of histones and transcription-related factors.* Microbiol Mol Biol Rev, 2000. **64**(2): p. 435-59.

91. Berndsen, C.E., et al., *Catalytic mechanism of a MYST family histone acetyltransferase.* Biochemistry, 2007. **46**(3): p. 623-9.

92. Liu, X., et al., *The structural basis of protein acetylation by the p300/CBP transcriptional coactivator.* Nature, 2008. **451**(7180): p. 846-50.

93. Lin, Y., et al., *Solution structure of the catalytic domain of GCN5 histone acetyltransferase bound to coenzyme A.* Nature, 1999. **400**(6739): p. 86-9.

94. Neuwald, A.F. and D. Landsman, *GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein.* Trends Biochem Sci, 1997. **22**(5): p. 154-5.

95. Yan, Y., et al., *Crystal structure of yeast Esa1 suggests a unified mechanism for catalysis and substrate binding by histone acetyltransferases.* Mol Cell, 2000. **6**(5): p. 1195-205.

96. Yang, X.J. and E. Seto, *HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention.* Oncogene, 2007. **26**(37): p. 5310-8.

97. Bouras, T., et al., *SIRT1 deacetylation and repression of p300 involves lysine residues 1020/1024 within the cell cycle regulatory domain 1.* J Biol Chem, 2005. **280**(11): p. 10264-76.

98. Black, J.C., et al., *The SIRT2 deacetylase regulates autoacetylation of p300.* Mol Cell, 2008. **32**(3): p. 449-55.

99. Han, Y., et al., *Acetylation of Sirt2 by p300 attenuates its deacetylase activity.* Biochem Biophys Res Commun, 2008. **375**(4): p. 576-80.

100. Marmorstein, R., *Structure and chemistry of the Sir2 family of NAD+-dependent histone/protein deactylases.* Biochem Soc Trans, 2004. **32**(Pt 6): p. 904-9.

101. Finnin, M.S., J.R. Donigian, and N.P. Pavletich, *Structure of the histone deacetylase SIRT2.* Nat Struct Biol, 2001. **8**(7): p. 621-5.

102. North, B.J. and E. Verdin, *Interphase nucleo-cytoplasmic shuttling and localization of SIRT2 during mitosis.* PLoS One, 2007. **2**(8): p. e784.

103. Kouzarides, T., *Acetylation: a regulatory modification to rival phosphorylation?* EMBO J, 2000. **19**(6): p. 1176-9.

104. Sakaguchi, K., et al., *DNA damage activates p53 through a phosphorylation-acetylation cascade.* Genes Dev, 1998. **12**(18): p. 2831-41.

105. Yang, X.J., *Multisite protein modification and intramolecular signaling.* Oncogene, 2005. **24**(10): p. 1653-62.

106. Thompson, P.R., et al., *Transcriptional coactivator protein p300. Kinetic characterization of its histone acetyltransferase activity.* J Biol Chem, 2001. **276**(36): p. 33721-9.

107. Yuan, L.W. and A. Giordano, *Acetyltransferase machinery conserved in p300/CBP-family proteins.* Oncogene, 2002. **21**(14): p. 2253-60.

108. Bedford, D.C., et al., *Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases.* Epigenetics, 2010. **5**(1): p. 9-15.

109. Radhakrishnan, I., et al., *Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions.* Cell, 1997. **91**(6): p. 741-52.

110. De Guzman, R.N., et al., *Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP.* J Mol Biol, 2000. **303**(2): p. 243-53.

111. De Guzman, R.N., et al., *CBP/p300 TAZ1 domain forms a structured scaffold for ligand binding.* Biochemistry, 2005. **44**(2): p. 490-7.

112. Legge, G.B., et al., *ZZ domain of CBP: an unusual zinc finger fold in a protein interaction module.* J Mol Biol, 2004. **343**(4): p. 1081-93.

113. Lin, C.H., et al., *A small domain of CBP/p300 binds diverse proteins: solution structure and functional studies.* Mol Cell, 2001. **8**(3): p. 581-90.

114. Kalkhoven, E., *CBP and p300: HATs for different occasions.* Biochem Pharmacol, 2004. **68**(6): p. 1145-55.

115. Mujtaba, S., et al., *Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation.* Mol Cell, 2004. **13**(2): p. 251-63.

116. Wei, L., et al., *Differential binding modes of the bromodomains of CREB-binding protein (CBP) and p300 with acetylated MyoD.* Biochem Biophys Res Commun, 2008. **368**(2): p. 279-84.

117. Kalkhoven, E., et al., *The PHD type zinc finger is an integral part of the CBP acetyltransferase domain.* Mol Cell Biol, 2002. **22**(7): p. 1961-70.

118. Bordoli, L., et al., *Functional analysis of the p300 acetyltransferase domain: the PHD finger of p300 but not of CBP is dispensable for enzymatic activity.* Nucleic Acids Res, 2001. **29**(21): p. 4462-71.

119. Chen, J., F.M. Ghazawi, and Q. Li, *Interplay of bromodomain and histone acetylation in the regulation of p300-dependent genes.* Epigenetics, 2010. **5**(6): p. 509-15.

120. Manning, E.T., et al., *p300 forms a stable, template-committed complex with chromatin: role for the bromodomain.* Mol Cell Biol, 2001. **21**(12): p. 3876-87.

121. Wang, L., et al., *Structure and chemistry of the p300/CBP and Rtt109 histone acetyltransferases: implications for histone acetyltransferase evolution and function.* Curr Opin Struct Biol, 2008. **18**(6): p. 741-7.

122. Janknecht, R., *The versatile functions of the transcriptional coactivators p300 and CBP and their roles in disease.* Histol Histopathol, 2002. **17**(2): p. 657-68.

123. Reynoird, N., et al., *Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains.* EMBO J, 2010. **29**(17): p. 2943-52.

124. Studier, F.W., *Protein production by auto-induction in high density shaking cultures.* Protein Expr Purif, 2005. **41**(1): p. 207-34.

125. Velazquez Campoy, A. and E. Freire, *ITC in the post-genomic era...? Priceless.* Biophys Chem, 2005. **115**(2-3): p. 115-24.

126. Mertens, H.D. and D.I. Svergun, *Structural characterization of proteins and complexes using small-angle X-ray solution scattering.* J Struct Biol, 2010. **172**(1): p. 128-41.

127. David, G., Pérez, J., *Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline.* J. Appl. Cryst., 2009. **42**(5): p. 892-900.

128. Incardona, M.F., et al., *EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis.* J Synchrotron Radiat, 2009. **16**(Pt 6): p. 872-9.

129. Kabsch, W., *Xds.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 125-32.

130. Kabsch, W., *Integration, scaling, space-group assignment and post-refinement.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 133-44.

131. Leslie, A.G., *The integration of macromolecular diffraction data.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 1): p. 48-57.

132. Evans, P.R., *An introduction to data reduction: space-group determination, scaling and intensity statistics.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 282-92.

133. McCoy, A.J., et al., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.

134. Drenth, J., *Principles of Protein X-Ray Crystallography (Third Edition)*, ed. S.A.T.i. Chemistry. 2011, Heidelberg (DE): Springer Verlag.

135. Rupp, B., *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology.* 2011: Garland Pub.

136. Li, W. and F. Li, *Cross-crystal averaging with search models to improve molecular replacement phases.* Structure, 2011. **19**(2): p. 155-61.

137. Mathew, E., A. Mirza, and N. Menhart, *Liquid-chromatography-coupled SAXS for accurate sizing of aggregating proteins.* J Synchrotron Radiat, 2004. **11**(Pt 4): p. 314-8.

138. Jin, Q., et al., *Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation.* EMBO J, 2011. **30**(2): p. 249-62.

139. Stavropoulos, P., et al., *Molecular basis for the autoregulation of the protein acetyl transferase Rtt109.* Proc Natl Acad Sci U S A, 2008. **105**(34): p. 12236-41.

140. Tang, Y., et al., *Fungal Rtt109 histone acetyltransferase is an unexpected structural homolog of metazoan p300/CBP.* Nat Struct Mol Biol, 2008. **15**(7): p. 738-45.

141. Lin, C. and Y.A. Yuan, *Structural insights into histone H3 lysine 56 acetylation by Rtt109.* Structure, 2008. **16**(10): p. 1503-10.

142. Albaugh, B.N., et al., *Autoacetylation of the histone acetyltransferase rtt109.* J Biol Chem, 2011. **286**(28): p. 24694-701.

143. Kadlec, J., et al., *Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1.* Nat Struct Mol Biol, 2011. **18**(2): p. 142-9.

144. Wang, Z., et al., *Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression.* Cell, 2010. **141**(7): p. 1183-94.

145. Dekker, F.J. and H.J. Haisma, *Histone acetyl transferases as emerging drug targets.* Drug Discov Today, 2009. **14**(19-20): p. 942-8.

APPENDIX

INTRODUCTION EN FRANÇAIS

Une collaboration, initiée avec le laboratoire de Saadi Khochbin à l'institut Albert Bonniot (IAB) de Grenoble, a conduit à une publication dans EMBO Journal (Septembre 2010). Ce laboratoire a étudié une protéine de fusion oncogénique appelée Brdt4-NUT résultant d'une translocation chromosomique qui fusionne le gène de NUT (*Nuclear protein in testis*) sur le chromosome 15 avec le gène de Brdt4 (une protéine contenant un double bromodomaine) sur le chromosome 19. La protéine de fusion Brdt4-NUT est impliquée dans l'apparition d'une forme grave de cancer appelée NUT *midline carcinoma* (NMC). Cette maladie est très rare et se contracte à l'âge moyen de 25 ans. Dès lors, l'espérance de vie moyenne du malade n'est que d'un an.

Le laboratoire de Saadi Khochbin a été en mesure de démontrer que le domaine NUT de la protéine de fusion peut se lier à p300 et ainsi l'activer. L'hyper-activation de p300 est obtenue par la formation d'une forte interaction entre le domaine CH3 de p300 (en particulier la région de TAZ2) et le domaine NUT. Aussi, Brdt4-NUT peut recruter p300 séquentiellement à travers le domaine NUT et il est capable de former des foyers de chromatine qui sont hyper-acétylés et transcriptionnellement inactifs. La séquestration du p300 au sein de ces foyers entraîne l'inactivation de la protéine anti-tumeur p53. Cela représente le principal mécanisme d'oncogenèse induit par la protéine de fusion conduisant à la transformation maligne des cellules.

Afin de démontrer que l'interaction entre p300 et NUT est directe, j'ai exprimé et purifié en cellules d'insectes p300 (résidus 324-2094) contenant le domaine CH3. En utilisant cette protéine, il a été démontré que l'interaction entre p300 et NUT était directe et qu'elle se faisait via le domaine CH3. De plus, la construction contenant le 'noyau' protéique de p300 n'est pas capable de former un lien avec NUT, sachant que le domaine CH3 n'y est pas présent. Enfin, des analyses d'activité ont été menées pour démontrer que l'activité catalytique de p300 est stimulée par son interaction avec NUT.

Cette étude représente l'une des preuves indiquant que le domaine CH3 de p300 peut jouer un rôle important dans la régulation de son activité enzymatique. Il introduit également l'idée que la séquestration de p300, résultant dans l'inactivation de p53, peut être utilisée comme un mécanisme d'oncogenèse par plusieurs protéines oncogéniques. Enfin, cette étude montre que l'activité, en dehors du contexte, d'un facteur spécifique testiculaire peut modifier sensiblement les fonctions cellulaires vitales et contribuer de manière significative à la transformation maligne des cellules.

Mots clés

Brdt, facteur de cancer testiculaire, oncogène, NMC, p300/CBP, acétylation, p53.

INTRODUCTION IN ENGLISH

During my thesis, we started a study in collaboration with Saadi Khochbin and his lab at the IAB (Institute Albert Bonniot, Grenoble). The Khochbin lab was working on an oncogenic fusion protein, Brd4-NUT, resulting from a chromosomal translocation which fuses the NUT gene (Nuclear protein in Testis) on chromosome 15 to the Brd4 gene (a double bromodomain containing protein) on chromosome 19. Brd4-NUT is involved in the appearance of a highly aggressive carcinoma called NUT midline carcinoma (NMC) which is very rare and has an average young age of appearance (25 years old) with a mean survival of one year.

The Khochbin lab could demonstrate that the NUT moiety of the fusion protein can interact with and activate p300. Hyper activation is accomplished by the formation of a stable binding between the p300 CH3 domain (in particular the TAZ2 region) and NUT. Brd4-NUT sequentially recruits p300 though its NUT moiety in order to form hyperacetylated and trascriptionally inactive chromatin foci. Sequestration of p300 into these foci leads to p53 inactivation, which is the principal oncogenic mechanism used by the fusion protein for malignant cell transformation.

To demonstrate that the interaction between p300 and NUT is direct, I have expressed and purified p300 (residues 324-2094) containing the CH3 domain from *Hi5* insect cells. Using this construct it was possible to show that the interaction between p300 and NUT is direct and mediated by the CH3 domain. Indeed, the p300 'core' construct lacking the CH3 region was unable to bind NUT. Moreover, HAT assays were performed to demonstrate that the HAT activity is stimulated by the binding of NUT.

The study reported here suggests that the CH3 domain of p300 may play an important role in protein regulation. Moreover, it introduces the idea that sequestration of p300 and consequent p53 inactivation can be used as a mechanism of oncogenesis by protein oncogenes. Finally, this study demonstrates how the off-context activity of a testis-specific factor can markedly alter vital cellular functions and significantly contribute to malignant cell transformation.

Key words

Brdt, cancer testis factor, oncogene, NMC, p300/CBP, acetylation, p53.

The EMBO Journal (2010) 29, 2943–2952 | © 2010 European Molecular Biology Organization | All Rights Reserved 0261-4189/10
www.embojournal.org

THE EMBO JOURNAL

Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains

Nicolas Reynoird[1], Brian E Schwartz[2], Manuela Delvecchio[3], Karin Sadoul[1], David Meyers[4,5], Chandrani Mukherjee[4,5], Cécile Caron[1], Hiroshi Kimura[6], Sophie Rousseaux[1], Philip A Cole[4,5], Daniel Panne[3], Christopher A French[2] and Saadi Khochbin[1,*]

[1]INSERM, U823, Université Joseph Fourier—Grenoble 1, Institut Albert Bonniot, Grenoble, France, [2]Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA, [3]EMBL Grenoble, Grenoble, France, [4]Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA, [5]Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD, USA and [6]Graduate School of Frontier Biosciences, Osaka University, Suita, Osaka, Japan

In a subset of poorly differentiated and highly aggressive carcinoma, a chromosomal translocation, t(15;19)(q13;p13), results in an in-frame fusion of the double bromodomain protein, BRD4, with a testis-specific protein of unknown function, NUT (nuclear protein in testis). In this study, we show that, after binding to acetylated chromatin through BRD4 bromodomains, the NUT moiety of the fusion protein strongly interacts with and recruits p300, stimulates its catalytic activity, initiating cycles of BRD4–NUT/p300 recruitment and creating transcriptionally inactive hyperacetylated chromatin domains. Using a patient-derived cell line, we show that p300 sequestration into the BRD4–NUT foci is the principal oncogenic mechanism leading to p53 inactivation. Knockdown of BRD4–NUT released p300 and restored p53-dependent regulatory mechanisms leading to cell differentiation and apoptosis. This study demonstrates how the off-context activity of a testis-specific factor could markedly alter vital cellular functions and significantly contribute to malignant cell transformation.

The EMBO Journal (2010) **29,** 2943–2952. doi:10.1038/emboj.2010.176; Published online 30 July 2010
Subject Categories: chromatin & transcription; molecular biology of disease
Keywords: Brdt; cancer testis factor; feed forward; H3K56; oncogene

Introduction

Malignant cell transformation is associated with a global disruption of genetic and epigenetic mechanisms leading to both aberrant gene activation and silencing. Although oncogenic gene silencing and activation and/or amplification of oncogenes are well documented, the impact of aberrant activation of normally silenced tissue-specific genes, that is, testis-specific genes, which is known to occur in many somatic cancers, is much less studied (Rousseaux and Khochbin, 2009).

The NUT midline carcinoma (NMC) refers to a group of malignant and highly lethal cancers, occurring in children and adults, which arise from chromosomal translocations systematically involving the *NUT* (NUclear protein in Testis) gene on chromosome 15q14. The function of *NUT* gene is unknown and it is normally expressed in testis (French, 2008). In the majority of NMC cases (two-third), the chromosomal translocation fuses *NUT* to the *BRD4* gene on chromosome 19 (French *et al*, 2003). A detailed analysis of the fusion transcript showed that *BRD4* exon 10b, normally used in a splice variant encoding a large BRD4 isoform, is fused to *NUT* exon 2 (French *et al*, 2008). The *BRD4* gene encodes a double bromodomain-containing protein belonging to a specific family of transcription/chromatin regulators known as BET (Bromodomain and Extra Terminal; Florence and Faller, 2001). In contrast with *NUT*, *BRD4* is ubiquitously expressed in somatic cells (French *et al*, 2003; Wu and Chiang, 2007).

Interestingly, in an additional subset of NMCs, which has recently been characterized, the chromosomal translocation fuses *NUT* to the *BRD3* gene on chromosome 9 (French *et al*, 2008). The *BRD3* gene is a paralogue of *BRD4* and also a member of the BET family function of which is less studied but, similar to BRD4, it preferentially associates with acetylated histones (LeRoy *et al*, 2008). This important finding points to the fusion of *NUT* with a double bromodomain-containing member of the BET family as an important determinant in the oncogenic activity of the fusion protein. Although the molecular basis of the oncogenic activity of *BRD–NUT* fusions remains largely unknown, it is clearly established that their downregulation in NMC cell lines induces squamous differentiation and arrested growth (French *et al*, 2008).

The data presented here explain why and how the fusion of *NUT* with genes encoding double bromodomain-containing factors of the BET family creates a new functional fusion protein, and shed light on a new oncogenic mechanism based on the off-context activity of a testis-specific factor.

Results

The BRD4–NUT fusion protein induces the formation of hyperacetylated chromatin domains

On its ectopic expression, BRD4–NUT forms discrete nuclear foci, which perfectly co-localize with hyperacetylated histone H3 and H4 chromatin domains (Figure 1A; Supplementary

*Corresponding author. INSERM, U823, Université Joseph Fourier—Grenoble 1, Insitut Albert Bonniot, Grenoble, F 38700, France. Tel.: +33 4 76 54 95 83; Fax: +33 4 76 54 95 95; E-mail: khochbin@ujf-grenoble.fr

Received: 15 February 2010; accepted: 6 July 2010; published online: 30 July 2010

©2010 European Molecular Biology Organization

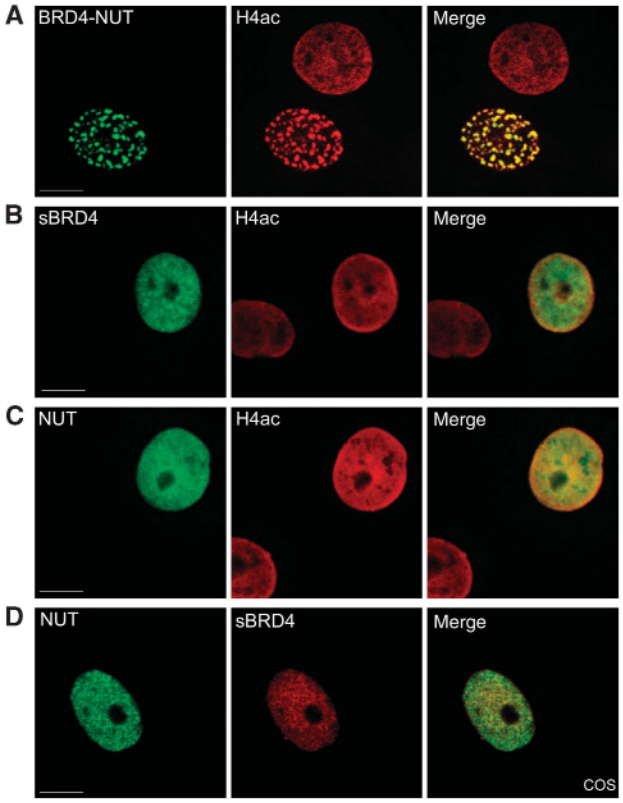


Figure 1 The BRD4–NUT fusion protein forms nuclear foci containing hyperacetylated chromatin domains. Cos7 cells were transfected with the indicated expression vectors, (**A**) GFP–BRD4–NUT, (**B**) HA–sBRD4, (**C**) HA–NUT and (**D**) HA–sBRD4 together with GFP–NUT. Ectopically expressed proteins were visualized by GFP or indirect fluorescence after anti-HA detection (**A**–**D**) and co-detected with an anti-pan acetylated histone H4 antibody (**A**–**C**). Bar: 10 μm.

Figure S1A and S2A). This particular pattern of nuclear localization is specific to the fusion protein because the expression of BRD4 alone (the short isoform of BRD4, approximately corresponding to the BRD4 part of the fusion protein, named here sBRD4) or that of NUT alone, or the co-expression of both (sBRD4 + NUT), does not induce the formation of hyperacetylated chromatin domains (Figure 1B–D, respectively). Moreover, the expression of the longer BRD4 (flBRD4) isoform with NUT did not allow the formation of distinct nuclear foci containing the two proteins or hyperacetylated chromatin (Supplementary Figure S1B and C).

It has previously been shown that BRD4 bromodomains interact with histone H4 acetylated at different positions and histone H3 acetylated on its lysine 14 (Dey *et al*, 2003; Lee and Chiang, 2009). As expected, these marks and other H4 and H3 acetylated forms were enriched in BRD4–NUT foci (Supplementary Figure S2A; data not shown). The absence of RNA polymerase II, or of its phosphorylated forms, suggests that these BRD4–NUT hyperacetylated foci are not associated with active gene transcription (Supplementary Figure S2B; also see Figure 2G). Consistent with this observation, no accumulation of H3K4me3 was observed in the BRD4–NUT foci (Supplementary Figure S2C, H3K4me3 panel). The absence of H3K9me3 in the foci shows that, despite the absence of RNA pol II, and in agreement with the presence of hyperacetylated histones, the BRD4–NUT foci are not of a heterochromatic nature (Supplementary Figure S2C, H3K9me3 panel).

The NUT moiety of BRD4–NUT specifically recruits CBP/p300

On the basis of these data, we hypothesized that BRD4–NUT could recruit cellular histone acetyltransferases (HATs), initiating cycles of BRD4–NUT/HAT recruitment and chromatin acetylation, leading to the formation of foci. Antibodies against several known HATs were used to detect their presence in the BRD4–NUT foci on its ectopic expression in Cos7 cells. Among the antibodies tested (P/CAF, Tip60, p300, CBP, HBO1 and HAT1), only anti-p300 and anti-CBP antibodies resulted in a clear accumulation of the endogenous protein in the BRD4–NUT foci (Figure 2A; Supplementary Figure S3; data not shown).

As no HAT had been reported to co-purify with BRD4 complexes (Jang *et al*, 2005; Yang *et al*, 2005), we assumed that NUT could be directing this efficient recruitment of cellular CBP/p300. This hypothesis was confirmed after immunoprecipitation of HA-tagged NUT, BRD4 and BRD4–NUT expressed in Cos7 cells. Endogenous p300 was observed to be associated only with NUT or BRD4–NUT (Figure 2B, lanes 2 and 4). In addition, we observed that NUT alone immunoprecipitates a strong associated HAT activity comparable with that of Gcn5 used as a control (Figure 2C, lanes 2 and 3).

To confirm the relevance of these observations based on the ectopic expression of our proteins of interest, we analysed the intranuclear localization of BRD4–NUT in a cell line derived from an aggressive, metastatic lung cancer arising in a young, non-smoking woman, HCC2429 (Haruki *et al*, 2005). Interestingly, the endogenous BRD4–NUT expressed in these cells (Figure 2D) also formed distinct nuclear foci very similar to the pattern observed after ectopic expression of BRD4–NUT (Figure 2E). In these cells, p300 and CBP were also observed to form foci, which perfectly localized with the BRD4–NUT nuclear domains (Figure 2E; data not shown).

Furthermore, we investigated the signature of p300 activity in the p300-containing foci in these cells. Histone H3K56 was very recently shown *in vivo* to be specifically acetylated by p300 (Das *et al*, 2009). Using a rabbit antibody specific for H3K56ac and a mouse monoclonal antibody against p300, we were able to clearly show a perfect co-localization of H3K56ac with p300 (Figure 2F). Another *in vivo* p300 mark is the acetylation of H3K18 (Ferrari *et al*, 2008; Horwitz *et al*, 2008). This modification also perfectly co-localized with p300 foci that co-localized with the BRD4–NUT foci (Figure 2F). Moreover, to definitely show the occurrence of hyperacetylated chromatin in the BRD4–NUT foci, several acetylated H4 and H3 sites, including H4K8ac, H3K14ac and H3K27ac, against which mouse monoclonal antibodies were available, were co-detected along with BRD4–NUT in the HCC2429 cells (Figure 2F; Supplementary Figure S2D).

Finally, to confirm that the endogenous BRD4–NUT foci are not active transcription sites, nascent RNAs labelled by BrU were detected both in HCC2429 and in a non-BRD4–NUT-expressing lung cancer cell line (A549). As shown in Figure 2G, none of the active transcription foci co-localize with the BRD4–NUT-containing domains, supporting the conclusion that the BRD4–NUT foci are transcriptionally silent chromatin domains.

©2010 European Molecular Biology Organization

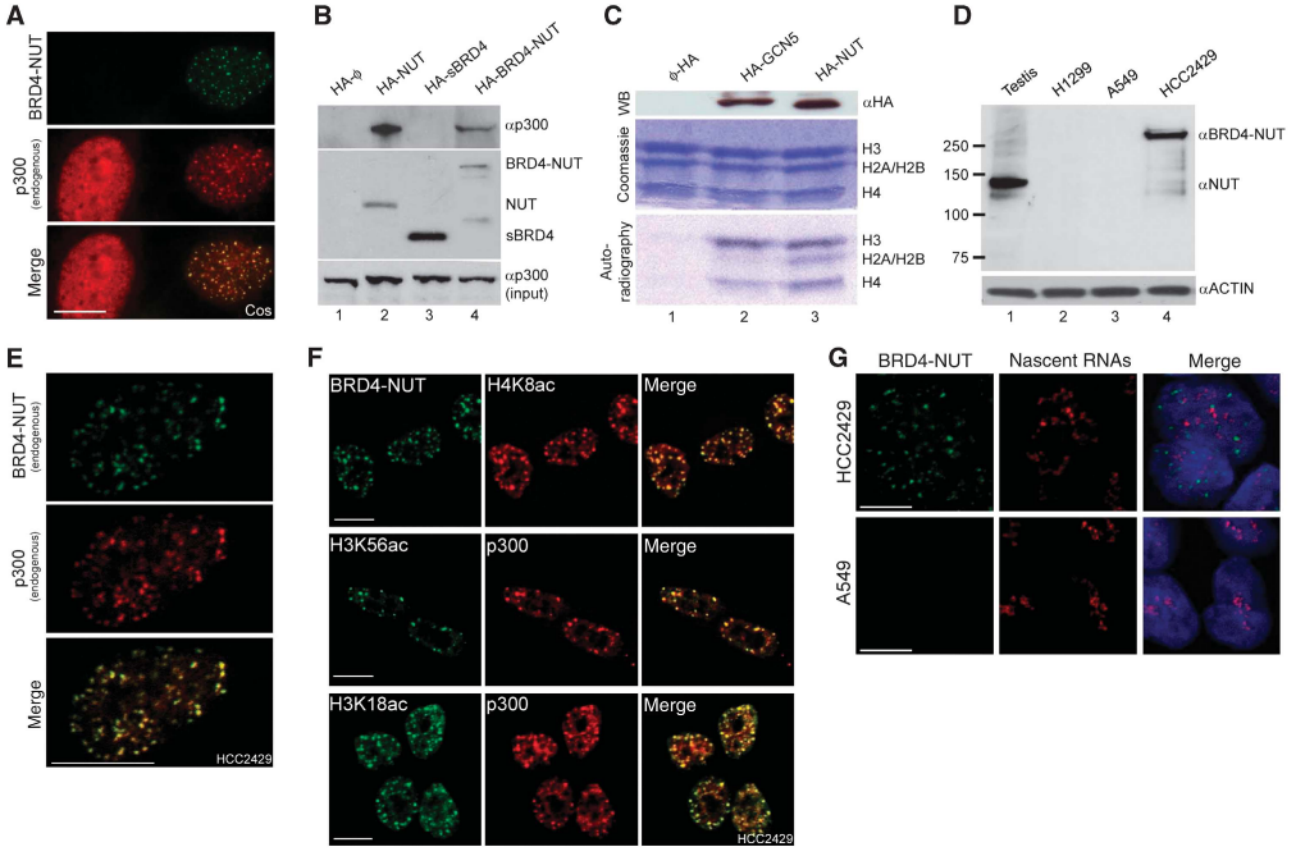


Figure 2 Specific recruitment of cellular p300 by the NUT moiety of BRD4–NUT to the nuclear foci. (**A**) Cos7 cells were transfected with GFP–BRD4–NUT and the endogenous cellular p300 was detected by immunofluorescence using an anti-p300 antibody. Both GFP and p300-related fluorescence were recorded and shown individually or together (merge panel). (**B**) Cos7 cells were transfected with the indicated HA-tagged vectors and, after anti-HA immunoprecipitation, HA-tagged proteins and associated endogenous p300 were visualized using the corresponding antibodies. The 'input' panel shows the amount of cellular p300 in each extract. (**C**) Cos7 cells were transfected with vectors expressing HA-tagged GCN5 or NUT and the corresponding empty vector (ϕ) and extracts were used to immunoprecipitate HA-tagged proteins. A fraction of immunoprecipitated materials was used to visualize proteins by an anti-HA western blot (anti-HA panel). Another fraction was used to monitor the HAT activity present in the immunoprecipitates using purified histones (lower panel, showing the corresponding autoradiography), which were visualized on a Coomassie-stained gel (middle panel). (**D**) Detection of endogenous BRD4–NUT with anti-NUT antibody in the HCC2429 cells compared with two other lung cancer cell lines (indicated). As a control, NUT was also detected in a rat total testis extract (lane testis). (**E**) Endogenous BRD4–NUT foci and p300 were detected in HCC2429 cells using the corresponding antibodies. (**F**) BRD4–NUT was co-detected with histone H4K8ac. The specific *in vivo* signatures of p300 activity, histones H3K56ac and H3K18ac, were also co-detected with endogenous p300 in HCC2429 cells (indicated). (**G**) BrU incorporation in nascent RNAs and BRD4–NUT foci were visualized in HCC2429 and A549 cells after a pulse labelling with BrUTP. Bar, 10 μm.

Domains involved in the direct interaction between NUT and p300

We next explored whether NUT and p300 interact directly and determined which domains are involved. First, a p300 fragment lacking the 870 N-terminal amino acids was observed to interact with NUT as efficiently as the full-length p300 (data not shown). This deletion mutant of p300 (Δ870) was used to further map the p300-interacting domains in NUT. An advantage of this p300 deletion mutant is that the interaction of both transfected p300 (Δ870) and endogenous p300 with NUT can be monitored. Figure 3A shows different fragments of both proteins used in these experiments. Co-immunoprecipitation experiments showed that full-length NUT, as well as a NUT deletion mutant containing only the N-terminal half of NUT (F1, see Figure 3A), interacted with both transfected Δ870 and endogenous p300 (Figure 3B, lanes 2 and 4). No interaction with p300 was observed when fragments of the C-terminal half of NUT were used (Figure 3B, lanes 6 and 8). We then tried to map the p300-interacting domain within the N-terminal half of NUT more precisely. Three fragments covering the N-terminal 593 amino acids of NUT were cloned (named F1a, b and c, respectively; Figure 3A) and tested as above. Only a limited region of NUT spanning amino acids 346–593 (F1c) interacted specifically with endogenous p300 (Figure 3C, lane 4). An analysis of the amino-acid sequence of NUT from different mammalian species showed that this region is highly conserved (data not shown), indicating its importance for the function of NUT. We further confirmed the interaction between the F1c domain of NUT and p300 by using a GST pull-down assay. GST alone or fused to the F1c fragment was incubated with Cos7 nuclear extracts. Figure 3D shows that the F1c fragment efficiently pulled down endogenous p300 (anti-p300 panel), and that this fraction contained *in vitro* HAT activity (autoradiography panel).

To map the NUT-interacting domain of p300, Cos7 cell extracts expressing different fragments of p300 were used in a pull-down assay with GST alone or the GST–NUT F1c fragment as described above. Figure 3E shows that the CH3 fragment of p300 encompasses the major binding site for

©2010 European Molecular Biology Organization

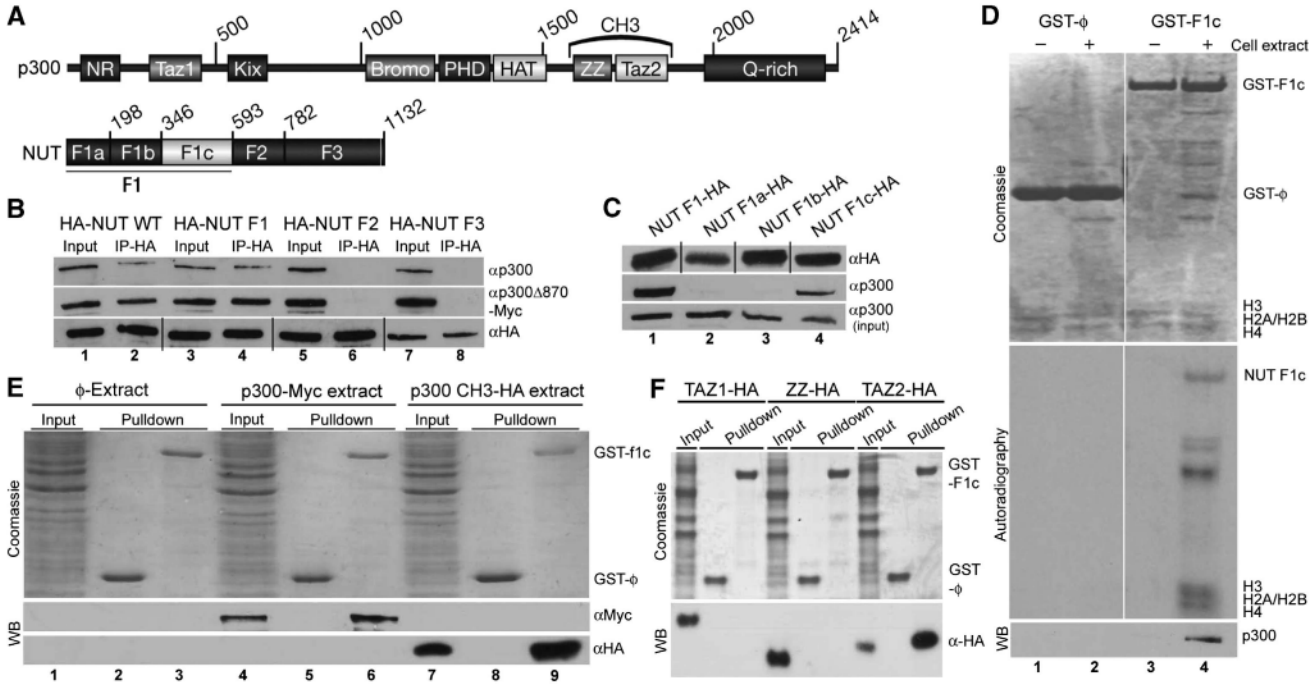


Figure 3 Mapping of domains in NUT and p300 involved in a direct interaction between the two proteins. (**A**) Schematic representation of domains and fragments used in both p300 and NUT. (**B**) Cos7 cells were transfected with Myc-tagged Δ870 p300 and the indicated fragments of HA-tagged NUT-expressing vectors and, after anti-HA immunoprecipitation, the immunoprecipitated proteins were visualized using the indicated antibodies. (**C**) The indicated HA-tagged NUT fragments were expressed in Cos7 cells and, after anti-HA immunoprecipitation (anti-HA panel), the association of the endogenous p300 was visualized (anti-p300 panel). The input panel indicates the presence of p300 in the different extracts. (**D**) Glutathione-bound GST or the GST–F1c fragment of NUT were incubated in a nuclear extract from Cos7 cells. Fractions of the bound proteins were either used to detect the presence of endogenous p300 (bottom panel) or to run a HAT assay using purified histones, as described previously. The Coomassie-stained gel (upper panel) shows GST and GST–F1c NUT before and after the pull down. The gel was then dried and used to visualize acetylated proteins after autoradiography (middle panel). (**E**) GST or GST–F1c were used to pull down either Myc-tagged p300 WT or the HA-tagged CH3 domain of p300. Extracts and proteins used in the pull down are shown after the Coomassie staining of the gel (upper panel). The proteins pulled down were visualized using the indicated antibodies. (**F**) In an experiment similar to (**E**), the interaction of GST or GST–F1c with the indicated fragments of p300 was monitored. Please note that as the HA-tagged proteins were of different sizes (**B**, **C**), only the corresponding areas of the films are shown and separated by lines. A full-colour version of this figure is available at *The EMBO Journal* Online.

NUT (lane 9). Two sub-domains have been identified in the CH3 domain and are known as ZZ and TAZ2. Pull-down experiments using extracts from cells expressing HA-tagged ZZ, TAZ2 and the p300 TAZ1 domain, showed that binding of NUT was specific to the TAZ2 domain (Figure 3F).

We next wanted to determine whether the interaction between p300 and NUT is direct. Bacterially expressed GST–NUT F1c fragment and two versions of p300 (called p300L for amino acids 324–2094 and p300S for 1045–1666 fragment, respectively with or without the CH3 domain) were purified after expression in a baculovirus system, and their interaction was tested in a GST pull-down assay. Purified p300 directly binds to NUT F1c (Figure 4A, lane 4), whereas no interaction was observed under the same conditions in the absence of the p300 CH3 domain (lane 3).

NUT strongly enhances p300 HAT activity

To investigate whether the NUT–p300 interaction could account for the hyperacetylation phenotype, we tested the effect of NUT on p300 HAT activity. Myc-tagged p300 with or without HA–NUT was transfected into Cos7 cells. Figure 4B shows that for the same amount of immunoprecipitated Myc–p300 (lanes 2 and 4, anti-Myc panel), the HAT activity was much stronger when p300 was co-expressed with NUT (autoradiography panel). We have also immunoprecipitated NUT using an anti-HA antibody. Figure 4B (anti-HA panel) shows that NUT efficiently co-immunoprecipitates p300 and that, here again, the p300 HAT activity is highly enhanced (compare lane 2, immunoprecipitated p300 without NUT, with lane 8, showing the same amount of p300 with NUT). Quantification of the labelled histones from several independent experiments by PhosphoImager and ImageQuant analysis showed that the presence of NUT increased p300 activity by about 8–10 times (data not shown).

We then took advantage of our purified system to directly test the stimulatory effect of NUT on p300 HAT activity. Increasing amounts of purified GST–NUT F1c fragment were incubated with the purified 324–2094 p300L fragment in an *in vitro* HAT assay. Histone H3 is a preferred target of p300 *in vivo* and was therefore used as a substrate in these assays. Figure 4C shows that purified F1c fragment clearly stimulates H3 acetylation by purified p300 in a dose-dependent manner. Using several modification-specific antibodies, we tested the stimulatory effect of the NUT F1c fragment on the specificity of p300 HAT activity. We observed that p300 alone acetylated all H3 acetyl-acceptor lysines except K56. Addition of GST–F1c enhanced the acetylation of H3K9, K14, K18 and K27 in a dose-dependent manner. Interestingly, in the presence of NUT F1c, p300 was able to acetylate H3K56, indicating that NUT might also specifically stimulate H3K56

 ©2010 European Molecular Biology Organization

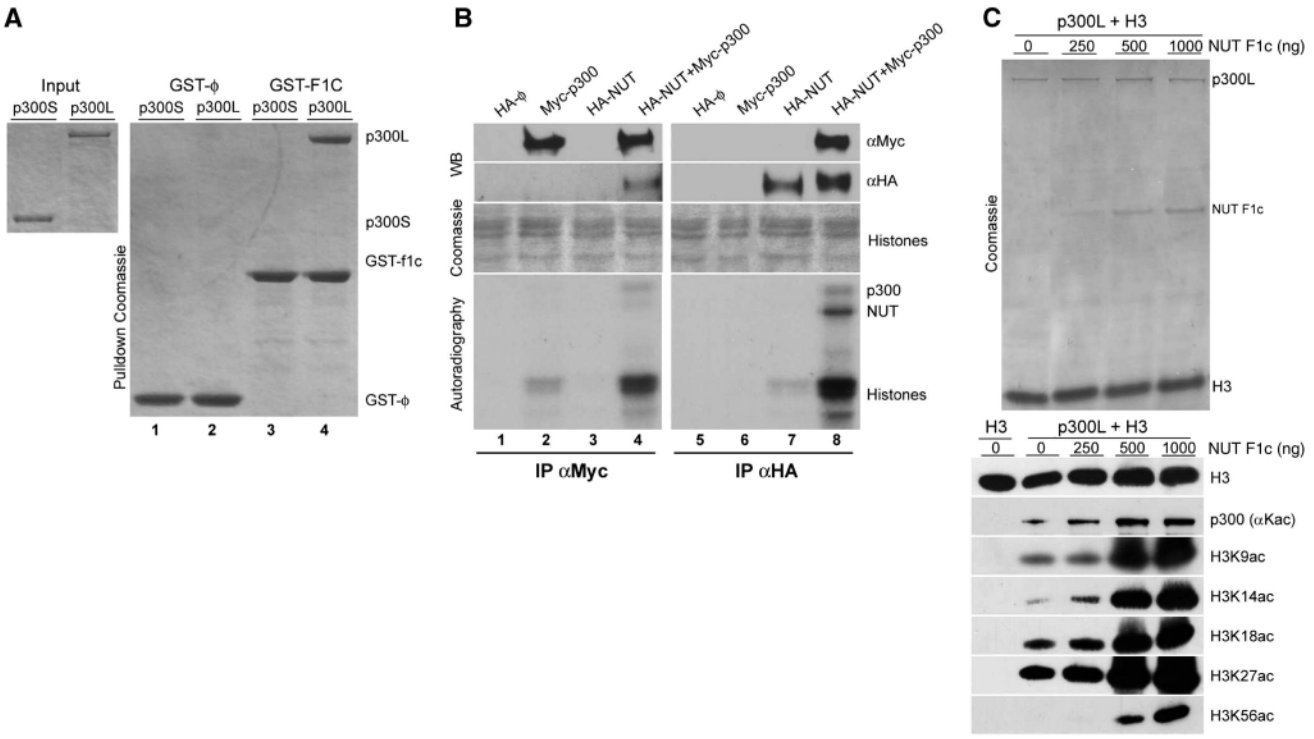


Figure 4 Direct enhancement of p300 HAT activity by NUT. (**A**) p300L (amino acids 324–2094 fragment) or p300S, a deletion mutant lacking its CH3 domain (amino acid 1045–1666 fragment) were expressed using a baculovirus system and purified. In parallel, GST and GST–F1c were expressed in bacteria and purified and used in a pull-down assay. The gel shows proteins present in the input (left panel) and after the pull down (right panel). (**B**) Cos7 cells were transfected with vectors expressing the indicated tagged proteins alone or together. After an immunoprecipitation using anti-Myc (p300) and anti-HA (NUT) antibodies, a fraction was used to monitor protein immunoprecipitations (upper panels). Another fraction was used for an *in vitro* HAT assay as described in Figure 2C. Coomassie-stained histones and the corresponding autoradiography are shown (middle and lower panels, respectively). (**C**) Purified p300 was incubated with histone H3 in the absence (0) or presence of increasing amounts of purified GST–F1c fragment of NUT (Coomassie panel). Aliquots corresponding to each incubation conditions were used to monitor site-specific histone H3 acetylation using antibodies recognizing the indicated acetylated sites. Anti-acetylated lysine and anti-H3 antibodies were used to show p300 autoacetylation and H3 loading, respectively. A full-colour version of this figure is available at *The EMBO Journal* Online.

acetylation. The same *in vitro* experiment was repeated using another fragment of NUT, F1b, incapable of interacting with p300 (Figure 3C). In contrast to F1c, this p300 non-interacting region of NUT had no effect on p300 HAT activity (Supplementary Figure S4).

BRD4–NUT mediates histone acetylation propagation through a feed-forward mechanism

The data obtained thus far suggested that the BRD4–NUT fusion could create a unique factor, binding acetylated chromatin through its bromodomains, and providing a platform for the recruitment of p300 and the stimulation of its HAT activity through NUT. An original feed-forward mechanism could then be initiated, through cycles of p300-mediated acetylation of adjacent nucleosomes and additional BRD4–NUT recruitments, leading to the formation of the observed foci. To test this hypothesis, we evaluated several of its testable aspects. First, BRD4–NUT foci formation and the corresponding chromatin acetylation should be abrogated by inactivating BRD4 bromodomains. Figure 5A shows that indeed, BRD4–NUT bearing inactivating mutations in its first bromodomain is unable to form foci and sustain local histone acetylation. Second, we reasoned that ectopic expression of p300 should stimulate the propagation of BRD4–NUT and hence lead to a decrease in the number of small BRD4–NUT foci in favour of the formation of larger foci. Figure 5B shows an example of two cells expressing BRD4–NUT, one of which also ectopically expresses p300. As predicted, fewer but larger BRD4–NUT foci are observed in the Myc–p300-expressing cells. Although p300 is known to form nuclear foci when expressed in cells (Eckner *et al*, 1994), p300 present in the BRD4–NUT foci, however, seemed to be of a different nature (Figure 5B, compare p300 in the BRD4–NUT foci with p300 alone shown in the inset). To quantify the effect of the ectopic expression of p300, the foci number and the area occupied by these foci were determined and compared with those of cells expressing only BRD4–NUT. This quantitative analysis shows that the ectopic expression of p300 leads to an approximately four-fold enlargement of BRD4–NUT foci mean area per cell, as well as a decrease in the total number of foci (Figure 5B, histograms).

We then used the HCC2429 carcinoma cell line and observed that the overexpression of p300-interacting region of NUT, F1c, or the NUT interaction region of p300, CH3, leads to the dispersion of p300 in the nucleus (Figure 5C). Notably, anti-NUT antibody recognizes the transfected F1c fragment, leading to a more intense NUT labelling in F1c-transfected cells.

We also anticipated that a continuous p300 activity should be required to keep BRD4–NUT concentrated at the foci and, consequently, that the inhibition of p300 catalytic activity should lead to the dispersion of BRD4–NUT foci. Accordingly,

©2010 European Molecular Biology Organization

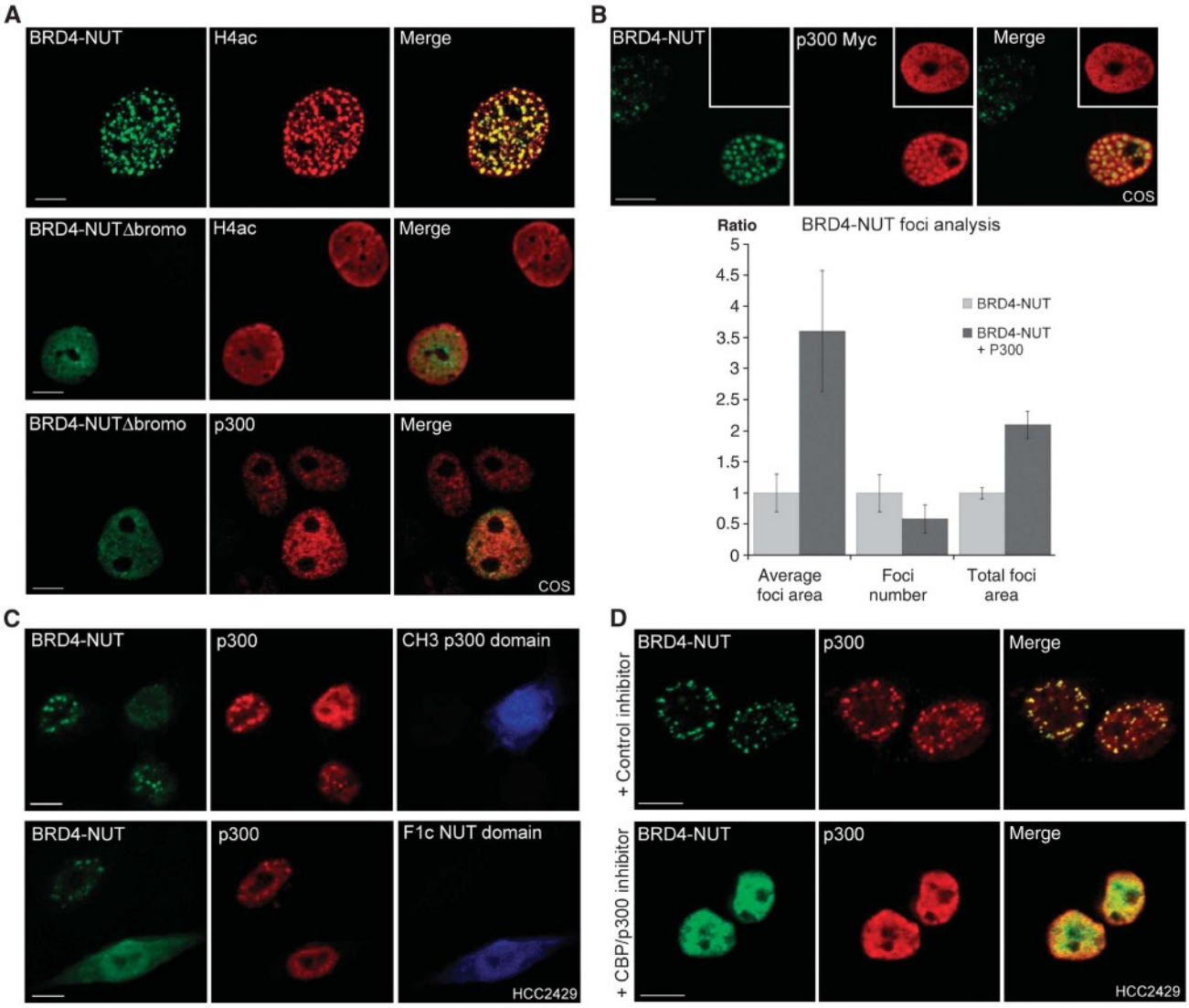


Figure 5 Propagation of chromatin acetylation through a feed-forward mechanism induced by BRD4–NUT. (**A**) The first BRD4 bromodomain was inactivated by site-directed mutagenesis (BRD4–NUTΔbr1) and the capacity of this mutant to form nuclear foci and acetylated chromatin domains enriched in p300 was analysed after the transfection of Cos7 cells. (**B**) Co-transfection of p300 together with GFP–BRD4–NUT leads to the enlargement of BRD4–NUT foci. The inset shows a cell expressing Myc–p300 alone. The histogram shows the results of quantitative evaluations (ratios of the values observed in p300-transfected cells to those observed in non-transfected cells) of the mean foci area (left), mean foci number (middle) and mean total foci area per cell (right), obtained as described in Materials and methods section. (**C**) HCC2429 were transfected either with the NUT interaction domain of p300, CH3 (upper panel), or the p300-interacting domain of NUT, F1c (lower panel), and the formation of foci by the endogenous p300 was monitored. NUT–F1c and p300–CH3 were detected with an anti-HA antibody, endogenous p300 with an anti-p300 antibody and BRD4–NUT with an anti-NUT antibody. Please note that the anti-NUT antibody also recognizes the NUT–F1c fragment. The images are representative of nearly 100% of the cells expressing either fragments. (**D**) HCC2429 cells were treated with the small p300 inhibitor, C646 (+ p300 inhibitor) or its inactive analogue (+ control inhibitor) and BRD4–NUT and p300 were visualized with the corresponding antibodies. The images are representative of nearly 100% of the cells. Bar: 10 μm.

a newly developed small molecule inhibitor of CBP/p300, C646 and its inactive analogue (Bowers *et al*, 2010), were used to test this hypothesis. Figure 5D shows that the treatment of cells with the specific CBP/p300 inhibitor, but not with its inactive analogue, leads to a clear dispersion of the BRD4–NUT and p300 foci.

Finally, our model predicts that the activity of cellular HDACs should oppose the p300-dependent propagation of BRD4–NUT foci. Accordingly, the inhibition of cellular HDACs by increasing the genome-wide histone acetylation should also lead to the propagation of BRD4–NUT and the dispersion of the foci. This was indeed the case because the treatment of HCC2429 with an HDAC inhibitor, TSA, induced complete dispersion of the foci (Supplementary Figure S5A). Besides, monitoring the amounts of BRD4–NUT after TSA treatment, we also noticed that this treatment induces a significant downregulation of BRD4–NUT at the time when an accumulation of the cell cycle regulator p21 was observed (Supplementary Figure S5B). As TSA treatment did not change the amount of BRD4–NUT encoding mRNA in HCC2429 cells (Supplementary Figure S5C), the BRD4–NUT downregulation probably occurred through a post-translational mechanism. This downregulation of BRD4–NUT was also observed after the treatment of cells with other unrelated

 ©2010 European Molecular Biology Organization

class I/II HDAC inhibitors, SAHA (vorinostat) and butyrate, but not with the class III HDAC inhibitor, nicotinamide, nor after a genotoxic treatment by etoposide (Supplementary Figure S5D).

Altogether these data are consistent with a mechanism in which the bromodomains and the NUT moiety of BRD4–NUT act synergistically as a potent histone hyperacetylator, action of which is counteracted by cellular HDACs, leading to the formation of hyperacetylated chromatin foci.

Inactivation of p53 regulatory circuits by BRD4–NUT

Our data show that BRD4–NUT foci titrate out cellular p300 into a limited number of transcriptionally silent hyperacetylated chromatin domains. Accordingly, we predicted that important p300-dependent cellular functions should also be altered in BRD4–NUT-expressing HCC2429 cells. CBP/p300 has an important role in mediating p53-dependent cellular responses by acetylating p53 (for review, see Grossman (2001)). We therefore hypothesized that p53 regulatory activity could be disrupted in the HCC2429 cells. Using an antibody specifically recognizing p53 acetylated by CBP/p300 (K373ac and K382ac), we observed that in HCC2429 cells acetylated p53 is also present in distinct nuclear foci, which disappear after the knockdown of BRD4–NUT (by two independent anti-NUT siRNAs), strongly suggesting that acetylated p53 remains bound to the BRD4–NUT/p300 foci (Supplementary Figure S6A). As both anti-acetylated p53 and anti-NUT antibodies were generated in rabbit, the co-detection of acetylated p53 and BRD4–NUT was not possible in HCC2429 cells; however, in a transfection-based assay using GFP–BRD4–NUT, acetylated p53 indeed accumulated in the BRD4–NUT foci, whereas total p53 was observed all over the nucleus (Supplementary Figure S7A).

On the basis of these observations, we predicted that p53 should be inactive in HCC2429 cells and therefore unable to ensure known p53-dependent cellular responses, such as *p21* induction after a genotoxic assault. To test the activity of p53 in HCC2429 cells, A549 cells expressing wild-type p53, used as a control, and HCC2429 cells, also expressing wild-type p53 (verified by sequencing; data not shown) were treated with etoposide and the activation of one p53 target gene, *p21* was monitored. Although an etoposide treatment efficiently induced p21 accumulation in A549 cells, it was not associated with a *p21* gene response in HCC2429 cells (Figure 6A), supporting the hypothesis that BRD4–NUT is capable of interfering with p53 functions in these cells. To confirm the inactivating role of BRD4–NUT on p53, we then transiently expressed BRD4–NUT in A549 cells and showed that it severely hindered *p21* activation, normally occurring in these cells after an etoposide treatment (Figure 6B). In addition, we made use of a *p21* promoter reporter system to monitor the p53 transactivator capacity in the absence or presence of increasing amounts of BRD4–NUT. After co-transfection of BRD4–NUT and p53, as expected, acetylated p53 was sequestered in BRD4–NUT foci (Supplementary Figure S7A) and increasing amounts of BRD4–NUT significantly reduced its activator function on the *p21* promoter (Supplementary Figure S7B).

Finally, to show the direct involvement of BRD4–NUT in p53 inactivation, we knocked down BRD4–NUT using two independent anti-NUT siRNAs and showed an efficient restoration of *p21* gene expression (Figure 6C) corresponding to the dispersion of acetylated p53 in the whole nucleus (Supplementary Figure S6A, siRNA BRD4–NUT panels). We could also demonstrate that the restoration of p53 activity was associated with a spontaneous HCC2429 cell apoptosis as judged by the accumulation of active caspase 3 (Figure 6C; Supplementary Figure S6B), as well as cell differentiation, as judged by the accumulation of epithelial differentiation marker E-cadherin (Figure 6C). Moreover, the release of p300 and acetylated p53 after BRD4–NUT knockdown in HCC2429 cells correlated with an enhanced apoptotic cell response after a genotoxic assault (Supplementary Figure S6B).

As some of the cellular events described here may occur in a p53-independent manner, we focussed our attention on p21 and showed that p21 accumulation induced by BRD4–NUT knockdown, was abolished if p53 was also knocked down (Figure 6D). Finally, to extend this study to other p53 target genes, in addition to *p21*, we also considered the expression of *PUMA* and *GADD45* (Harris and Levine, 2005). Figure 6E shows that BRD4–NUT downregulation after the treatment of cells with two independent anti-NUT siRNAs leads to a significant accumulation of mRNA encoded by the three p53 target genes. Interestingly, the treatment of cells with a CBP/p300 inhibitor severely interfered with the induction of these genes, showing that CBP/p300 activity is required to activate p53 target genes after BRD4–NUT knockdown. Accordingly, a ChIP approach showed that the knockdown of BRD4–NUT leads to the recruitment of acetylated p53 and p300 on p21 promoter concomitant with *p21* gene activation (Figure 6F).

Altogether, these data clearly show a role for BRD4–NUT in the extinction of critical CBP/p300 functions through the sequestration of these enzymes in HCC2429 cells.

Discussion

The testis-specific factor, NUT, is central to the oncogenic activity of the fusion protein in NMCs because it is invariably involved in the resulting oncogenic chromosomal translocations (French, 2008). In addition, the fusion with a double bromodomain protein of the BET family, either BRD4 or BRD3, seems to constitute the second critical element in the pathological activity of the fusion protein (French *et al*, 2008). Here, the functional dissection of the BRD4–NUT protein reveals unique properties for this fusion protein and explains why selection operates towards the fusion of *NUT* with members of the BET family.

The fact that the fusion of NUT with BRD4 or BRD3, probably with distinct cellular functions (Wu and Chiang, 2007; LeRoy *et al*, 2008; Chiang, 2009), results in the same phenotype strongly suggests that their important functional contribution is related to the presence of their bromodomains that show similar acetylated histone-binding activities *in vivo* (Dey *et al*, 2003; LeRoy *et al*, 2008).

The protein, NUT, is a testis-specific protein of totally unknown function. On the basis of the data presented in this study, it can be speculated that the physiological role of NUT is related to its association with, and stimulation of p300 and/or other yet unknown HATs. This activity of NUT could contribute to the wave of genome-wide post-meiotic histone hyperacetylation that occurs in elongating spermatids just before their replacement by transition proteins (Boussouar

©2010 European Molecular Biology Organization

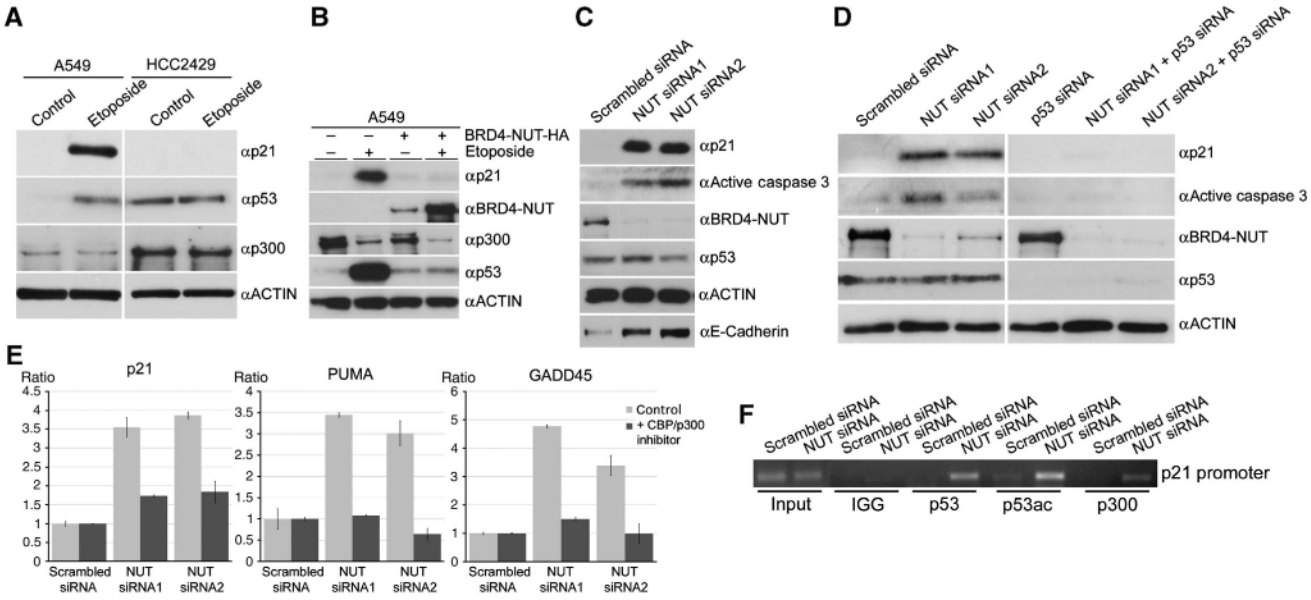


Figure 6 p53 inactivation in BRD4–NUT-expressing cells. (**A**) HCC2429 and A549 cells were treated with 20 μM etoposide for 15 h and p21, p53 and p300 were visualized as indicated. (**B**) A546 cells were transfected with an empty vector (−) or with a BRD4–NUT-expressing vector (+) and treated or not with etoposide as above (+ and −, respectively) and the indicated proteins were detected with the corresponding antibodies. (**C**) HCC2429 cells were treated with scrambled siRNAs or two specific anti-BRD4–NUT siRNAs and the indicated proteins were visualized as above. (**D**) HCC2429 cells were treated with the anti-BRD4–NUT siRNAs indicated above as well as with an anti-p53 siRNA and the indicated proteins were visualized using the corresponding antibodies. (**E**) The expression of *p21*, *PUMA* and *GADD45* were monitored after the knockdown of BRD4–NUT with the two above-mentioned siRNAs in the presence (dark grey histograms) or absence (light grey histograms) of p300 inhibitor. Data represent the mean ± s.d. values of at least three independent experiments. (**F**) The association to *p21* promoter of p53, acetylated p53 (p53ac) and p300 was monitored in HCC2429 cells before and after the knockdown of BRD4–NUT by siRNA as above. The PCR amplification shown here is representative of three independent ChIPs.

et al, 2008; Rousseaux *et al*, 2008) in which both NUT and p300 are expressed (data not shown).

Interestingly, the BRD4–NUT-dependent histone hyperacetylation evidenced here is not associated with transcription. This situation is similar to that observed in maturing post-meiotic male germ cells in which, despite histone hyperacetylation, there is a major turn-off in transcriptional activity associated with chromatin compaction. In BRD4–NUT-expressing cells, the absence of transcription, despite p300 recruitment and histone hyperacetylation, could also be associated with an acetylation-dependent compaction of chromatin in the foci. Indeed, an acetylation-dependent chromatin compaction was previously reported under the action of the testis-specific member of the BET family, Brdt (Pivot-Pajot *et al*, 2003; Govin *et al*, 2006; Moriniere *et al*, 2009). It is therefore possible that BRD4–NUT locally mimics a 'Brdt-like' activity in compacting acetylated chromatin domains, which would hinder the access of RNA Pol II to genes.

The dissection of the mechanisms underlying the formation of acetylated BRD4–NUT chromatin foci revealed a unique mechanism leading to the creation of large acetylated chromatin domains. Indeed, a BRD4–NUT mutant with an inactive bromodomain is incapable of forming BRD4–NUT acetylated chromatin domains, suggesting that the initial targeting of acetylated chromatin by the fusion protein is necessary to initiate the process. Once BRD4–NUT is in place, the recruitment of p300 and the stimulation of its HAT activity would allow the propagation step. As shown also for MAML1–p300 interaction, engagement of the CH3 domain by various protein ligands may be a general mechanism for enhancing the catalytic activity of p300 (Hansson *et al*, 2009). The extent of this propagation is limited by HDAC activity, which, by opposing histone acetylation, could slow down and eventually stop the propagation, thereby creating discrete foci. The presence and activity of p300 are critical for the maintenance of BRD4–NUT foci. Indeed, the overexpression of the p300-interacting domain of NUT, F1c, leads to the displacement of p300 and the dispersion of the BRD4–NUT foci. This experiment also reveals an important role for the CH3 domain of p300 in the maintenance of these foci. Interestingly, this domain of p300 is critical for the 'physiological' p300 foci formation because, according to early investigations, the CH3-interacting viral protein E1A also disrupts these foci (Lill *et al*, 1997). In addition, we show in this study that the continuous activity of p300 is also required for the stability of the BRD4–NUT foci, as its inhibition with a specific small molecule inhibitor disrupts the foci probably due to a dominant HDAC activity. Indeed, in patient-derived HCC2429 cells, the genome-wide histone hyperacetylation induced by TSA and the subsequent downregulation of BRD4–NUT leads to the dispersion of BRD4–NUT foci.

In the absence of TSA treatment, the high affinity binding of p300 to the NUT moiety of the fusion protein leads to the accumulation of the majority of cellular p300 into these transcriptionally inactive BRD4–NUT foci. On the basis of these findings, we also predicted that in BRD4–NUT-expressing cells important cellular functions depending on p300 should be altered. One of the functions of p300 is the control of p53 regulatory activity (Grossman, 2001). Indeed, we show in this study that the release of p300 after BRD4–NUT

 ©2010 European Molecular Biology Organization

downregulation in HCC2429 cells leads to the activation of p53 target genes in a CBP/p300-dependent manner. These studies allowed us to clearly demonstrate one of the oncogenic activities of the BRD4–NUT fusion protein. However, as CBP/p300 and BRD4 are involved in various critical cellular processes, the sequestration of these proteins in limited numbers of nuclear domains certainly affects other cellular functions hindering apoptotic cell response and cell differentiation. For instance, a deficiency of BRD4, which has important roles in polII transcription (Wu and Chiang, 2007), in BRD4–NUT-expressing cells could also be oncogenic. Indeed, it has been shown that $Brd4^{+/-}$ cells present increased chromosomal missegregation (Nishiyama *et al*, 2006).

The most important aspect of this study, however, is that it uncovers a new concept in oncogenesis. Indeed, we show in this study how the off-context activity of a testis-specific factor in a somatic cell becomes oncogenic. As the aberrant expression of testis-specific genes has been reported in many cancer cells, we propose that, at least in some cases, that is NUT, the product of these genes can be major contributors to malignant cell transformation (Rousseaux and Khochbin, 2009). This concept opens a new area of research in cancer biology.

Materials and methods

Plasmids

The BRD4–NUT GFP construction was kindly provided by Dr Dang, and then sub-cloned in a pcDNA 3.1 HA vector. p300 WT and ΔN870 Myc plasmids were generous gifts from Dr H Stunnenberg. The sBRD4 and NUT constructions were obtained after PCR amplification of their corresponding coding sequences from a human testis cDNA bank (Clontech) and cloned in pcDNA 3.1 HA and pEGFP-C vectors. Different NUT fragments (f1, f1a, f1b, f1c, f2 and f3) were amplified using the appropriate sets of primers and cloned in pcDNA 3.1 HA and in pGEX5 GST vectors. p300 CH3, TAZ1, ZZ and TAZ2 fragments were amplified from the p300WT plasmid and cloned into the pcDNA 3.1 HA vector. BRD4–NUT ΔBromo1 construct was obtained by deleting six critical amino acids of bromodomain 1 (81-PFQQPV-88) directly in BRD4–NUT pcDNA HA plasmid, using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). Mutations were then confirmed by DNA sequencing.

p53 Flag and p21 promoter luciferase plasmids were purchased from Addgene Organisation (www.addgene.org).

Cell culture, treatments, transfection assays, antibodies and gene reporter assay

Cos7 cells were maintained in DMEM medium, whereas H1299, A549 and HCC2429 cells were maintained in RPMI medium complemented with 10% SVF, 2% glutamine and 1% P/S. When needed, cells were treated with TSA 100 ng/ml (Sigma) for different times or with etoposide 20 μg/ml (Sigma).

p300 inhibitor, C646, or its inactive analogue (Bowers *et al*, 2010) was added to the culture medium from 1 h to complete siRNA incubation time (24–48 h), at 20 μg/ml.

Cells were seeded and grown to 60–70% confluence on the day of transfection. Transfections were performed with Lipofectamine 2000 (Invitrogen), using a ratio of 1 μl of Lipofectamine for 1 μg total quantity of vectors. The BRD4–NUT siRNA1 (forward: 5′-gcau cuaaugugaagacca-3′), BRD4–NUT siRNA2 (forward: 5′-gggauugcaga aaggacaa-3′) and scrambled siRNA were purchased from Eurogentec, whereas p53 siRNA was obtained from Qiagen (SI02655170, Hs_TP53_9). Cells were transfected with Lipofectamine RNAimax (Invitrogen) using the company instructions and cells were collected 24–48 h before analysis.

The primary antibodies used were: anti-NUT (Cell Signaling Technology), anti-p300 rabbit (N-15, Santa Cruz Biotechnology), anti-p300 mouse (05–257, Millipore), anti-BrdU (A21300, Invitrogen), anti-HA rat (for IP and IF, Roche), anti-HA mouse (for WB, Covance), anti-FlagM2 (F3165, Sigma), anti-Myc rabbit (ab9106, Abcam), anti-β-Actin (A5441 Sigma), anti-Histone H4 pan rabbit (05–858, Millipore), anti-H3 rabbit (05–928, Millipore), anti-H3K9ac rabbit (07–352, Millipore), anti-H3K14ac rabbit (07–353, Millipore), anti-H3K18ac rabbit (Millipore), anti-H3K56ac rabbit (2134–1, Epitomics), anti-RNA pol II (8WG16, H5, H14, Covance), anti p53 (DO7, Dako), anti-p53 K373ac-K382ac (06–758, Millipore), anti-p21/WAF1 (ab-1OP64, Calbiochem), anti-PARP (Boehringer), anti-cleaved-caspase 3 (9661S, Cell Signaling Technology), anti-E-cadherin (18–0223, Zymed Laboratories) and mouse anti-acetylated histones from H Kimura.

Gene reporter assay was carried out as detailed in the Supplementary data.

Immunofluorescence, BrUTP incorporation, immunoprecipitation, HAT assay, GST pull down, qRT–PCR and ChIP

Cells were grown on two- or four-well glass Lab-Tek (Nunc), transfected with 1 μg plasmid of interest, fixed in 4% paraformaldehyde for 5 min at room temperature and then permeabilized in 4% paraformaldehyde–0.1% Triton X-100 for 1 min at room temperature. Fixed and permeabilized cells were incubated in blocking buffer (PBS containing 5% skim milk) for 30 min at room temperature, then in primary antibody (except for GFP-tagged proteins) for 1 h at 37°C, then washed with PBS and incubated in the secondary antibody for 30 min at 37°C. After three washes in PBS, the cells were counterstained with Hoechst (250 ng/ml) and examined under a confocal microscope.

In situ transcription was monitored by BrUTP (Sigma) incorporation as described previously (Roussel *et al*, 1996) in conditions set up to reveal RNA pol I and RNA pol II transcription (Moore and Ringertz, 1973). Briefly, cells were grown on two- or four-well glass Lab-Tek (Nunc) and slightly fixed and permeabilized for 5 min at 4°C, then incubated for 30 min at 37°C with transcription buffer (100 mM Tris, 4 mM $MgCl_2$, 2 mM $MnCl_2$, 100 mM $(NH4)_2SO_4$, 0.1% B-mercaptoethanol, 150 mM sucrose, 0.6 mM ATP, GTP and CTP and 0.12 mM BrUTP). Incorporated BrU was then detected with 1/100 anti-BrdU antibody (Invitrogen) incubated for 1 h at 37°C.

For qRT–PCR, RNA was extracted from treated or non-treated cells using TRIzol reagent (Invitrogen) and reverse transcribed into cDNA with SuperScript III First-Strand Synthesis mix (Invitrogen). Human p21 (F: 5-CCTGTCACTGTCTTGTACCCT-3, R: 5-TTAGCAG CGGAACAAGGAGT-3), PUMA (F: 5-GACCTCAACGCACAGTACGA-3, R: 5-GAGATTGTACAGGACCCTCCA-3), GADD45a (F: 5-TTTGCAATA TGACTTTGGAGGA-3 R: 5-CATCCCCACCTTATCCAT-3) and BRD4–NUT (F: 5-AGGCTGTCATTCACCCTCAA-3, R: 5-CTCTTCCAAGGCCA TGAGTC-3) expressions were then monitored using Brilliant SYBR Green Master mix and MX3005P (Stratagene) and normalized by ACTIN or GAPDH expression.

ChIP experiment was carried out on the basis of X-ChIP protocol from Abcam with modifications. Briefly, 10^7 cells after siRNA treatment were cross-linked for 10 min at RT with 1% formaldehyde on a shaking platform and final 125 mM glycine was added to stop the reaction. Cells were then collected, washed twice with cold PBS and incubated for 10 min with cell lysis buffer (20 mM Tris–HCl (pH 8.0), 85 mM KCl and 0.5% NP40). Nuclei were re-suspended in Nuclei lysis buffer (50 mM Tris–HCl (pH 8.0), 10 mM EDTA and 1% SDS) and then sonicated to obtain 500-bp fragments of DNA. After centrifugation, the supernatant was diluted five times in ChIP buffer (20 mM Tris–HCl (pH 8.0), 1.1 mM EDTA, 0.01% SDS, 1.1% Triton X-100 and 167 mM NaCl) and incubated overnight at 4°C with appropriate antibodies, and then for 1 h with protein A dynabeads (Invitrogen). After intensive washes, twice in low-salt buffer (20 mM Tris–HCl (pH 8.0), 2 mM EDTA, 0.1% SDS, 1% Triton X-100 and 150 mM NaCl), twice in high-salt buffer (20 mM Tris–HCl (pH 8.0), 2 mM EDTA, 0.1% SDS, 1% Triton X-100 and 500 mM NaCl), once in LiCl buffer (0.25 M LiCl, 1% NP40, 20 mM Tris–HCl (pH 8.0), 1 mM EDTA and 1% deoxycholate) and finally in TE buffer (10 mM Tris–HCl (pH 8.0) and 1 mM EDTA), the DNA was eluted by incubating the beads for 30 min at RT with the elution buffer (1% SDS and 100 mM $NaHCO_3$). The DNA was then purified by phenol:chloroform and qRT–PCR was performed using primers for the p53-binding site on *p21* promoter (F: 5-GCTTGGGCAGCAG GCTG-3, R: 5-AGCCCTGTCGCAAGGATCC-3).

Detailed protocols for immunoprecipitation, HAT assay and GST pull down, as well as for insect cell culture, baculovirus cultivation and protein purifications are provided in the Supplementary data.

©2010 European Molecular Biology Organization

Quantitative analysis of BRD4–NUT Foci

Confocal microscopy images from different nuclei along the *z*-axis were analysed using the Metamorph software. The number of foci was counted in 50 nuclei and normalized with respect to the total nucleus area obtained using Metamorph software functions. The mean foci area was obtained by dividing the total number of BRD4–NUT foci areas in a given nucleus obtained by the Metamorph software as above, by the number of foci counted within the same nucleus.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank Martin Sos (Thomas Laboratory), Max-Planck Institute for Neurological Research and Cancer Genomics for providing us with the HCC2429 cell line. SK team also acknowledges the precious help of Sandrine Curtet in cell culture. The work in SK laboratory was supported by INCa-DHOS, ANR blanche 'EpiSperm' and 'Empreinte' and ARC research programmes. The work in CAF laboratory was supported by a Dana Farber/Harvard Cancer Center Nodal Award 5P30CA06516-44, US National Institutes of Health grant 1R01CA124633 and funds from the National Cancer Institute's Initiative for Chemical Genetics (Contract No. N01-CO-12400). DM, CM and PAC thank the NIGMS (GM62437) and FAMRI foundation for support. NR is a recipient of Rhone-Alpes region Ph D programme and of 'Fondation pour la Recherche Medicale'.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Boussouar F, Rousseaux S, Khochbin S (2008) A new insight into male genome reprogramming by histone variants and histone code. *Cell Cycle* **7:** 3499–3502

Bowers E, Yan G, Mukherjee C, Orry A, Wang L, Holbert M, Crump N, Hazzalin C, Liszczak G, Yuan H, Larocca C, Saldanha S, Abagyan R, Sun Y, Meyers D, Marmorstein R, Mahadevan L, Alani R, Cole P (2010) Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. *Chem Biol* **17:** 471–482

Chiang CM (2009) Brd4 engagement from chromatin targeting to transcriptional regulation: selective contact with acetylated histone H3 and H4. *F1000 Biol Rep* **1**, pii 98

Das C, Lucia MS, Hansen KC, Tyler JK (2009) CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature* **459:** 113–117

Dey A, Chitsaz F, Abbasi A, Misteli T, Ozato K (2003) The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc Natl Acad Sci USA* **100:** 8758–8763

Eckner R, Ewen ME, Newsome D, Gerdes M, DeCaprio JA, Lawrence JB, Livingston DM (1994) Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes Dev* **8:** 869–884

Ferrari R, Pellegrini M, Horwitz GA, Xie W, Berk AJ, Kurdistani SK (2008) Epigenetic reprogramming by adenovirus e1a. *Science* **321:** 1086–1088

Florence B, Faller DV (2001) You bet-cha: a novel family of transcriptional regulators. *Front Biosci* **6:** D1008–D1018

French CA (2008) Molecular pathology of NUT midline carcinomas. *J Clin Pathol* **63:** 492–496

French CA, Miyoshi I, Kubonishi I, Grier HE, Perez-Atayde AR, Fletcher JA (2003) BRD4–NUT fusion oncogene: a novel mechanism in aggressive carcinoma. *Cancer Res* **63:** 304–307

French CA, Ramirez CL, Kolmakova J, Hickman TT, Cameron MJ, Thyne ME, Kutok JL, Toretsky JA, Tadavarthy AK, Kees UR, Fletcher JA, Aster JC (2008) BRD–NUT oncoproteins: a family of closely related nuclear proteins that block epithelial differentiation and maintain the growth of carcinoma cells. *Oncogene* **27:** 2237–2242

Govin J, Lestrat C, Caron C, Pivot-Pajot C, Rousseaux S, Khochbin S (2006) Histone acetylation-mediated chromatin compaction during mouse spermatogenesis. *Ernst Schering Res Found Workshop* **57:** 155–172

Grossman SR (2001) p300/CBP/p53 interaction and regulation of the p53 response. *Eur J Biochem* **268:** 2773–2778

Hansson ML, Popko-Scibor AE, Saint Just Ribeiro M, Dancy BM, Lindberg MJ, Cole PA, Wallberg AE (2009) The transcriptional coactivator MAML1 regulates p300 autoacetylation and HAT activity. *Nucleic Acids Res* **37:** 2996–3006

Harris SL, Levine AJ (2005) The p53 pathway: positive and negative feedback loops. *Oncogene* **24:** 2899–2908

Haruki N, Kawaguchi KS, Eichenberger S, Massion PP, Gonzalez A, Gazdar AF, Minna JD, Carbone DP, Dang TP (2005) Cloned fusion product from a rare t(15;19)(q13.2;p13.1) inhibit S phase *in vitro*. *J Med Genet* **42:** 558–564

Horwitz GA, Zhang K, McBrian MA, Grunstein M, Kurdistani SK, Berk AJ (2008) Adenovirus small e1a alters global patterns of histone modification. *Science* **321:** 1084–1085

Jang MK, Mochizuki K, Zhou M, Jeong HS, Brady JN, Ozato K (2005) The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* **19:** 523–534

Lee AY, Chiang CM (2009) Chromatin adaptor Brd4 modulates E2 transcription activity and protein stability. *J Biol Chem* **284:** 2778–2786

LeRoy G, Rickards B, Flint SJ (2008) The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Mol Cell* **30:** 51–60

Lill NL, Grossman SR, Ginsberg D, DeCaprio J, Livingston DM (1997) Binding and modulation of p53 by p300/CBP coactivators. *Nature* **387:** 823–827

Moore GP, Ringertz NR (1973) Localization of DNA-dependent RNA polymerase activities in fixed human fibroblasts by autoradiography. *Exp Cell Res* **76:** 223–228

Moriniere J, Rousseaux S, Steuerwald U, Soler-Lopez M, Curtet S, Vitte AL, Govin J, Gaucher J, Sadoul K, Hart DJ, Krijgsveld J, Khochbin S, Muller CW, Petosa C (2009) Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* **461:** 664–668

Nishiyama A, Dey A, Miyazaki J, Ozato K (2006) Brd4 is required for recovery from antimicrotubule drug-induced mitotic arrest: preservation of acetylated chromatin. *Mol Biol Cell* **17:** 814–823

Pivot-Pajot C, Caron C, Govin J, Vion A, Rousseaux S, Khochbin S (2003) Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Mol Cell Biol* **23:** 5354–5365

Rousseaux S, Khochbin S (2009) New hypothesis for large-scale epigenome alterations in somatic cancer cells: a role for male germ-cell specific regulators. *Epigenomics* **1:** 153–161

Rousseaux S, Reynoird N, Escoffier E, Thevenon J, Caron C, Khochbin S (2008) Epigenetic reprogramming of the male genome during gametogenesis and in the zygote. *Reprod Biomed Online* **16:** 492–503

Roussel P, Andre C, Comai L, Hernandez-Verdun D (1996) The rDNA transcription machinery is assembled during mitosis in active NORs and absent in inactive NORs. *J Cell Biol* **133:** 235–246

Wu SY, Chiang CM (2007) The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *J Biol Chem* **282:** 13141–13145

Yang Z, Yik JH, Chen R, He N, Jang MK, Ozato K, Zhou Q (2005) Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* **19:** 535–545

©2010 European Molecular Biology Organization

Supplementary materials

Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains

Nicolas Reynoird[1], Brian E. Schwartz[2], Manuela Delvecchio[3], Karin Sadoul[1], David Meyers[4], Chandrani Mukherjee[4], Cécile Caron[1], Hiroshi Kimura[5], Sophie Rousseaux[1], Philip A. Cole[4], Daniel Panne[3], Christopher A. French[2], Saadi Khochbin[1*]

1 - INSERM, U823; Université Joseph Fourier - Grenoble 1; Institut Albert Bonniot, Grenoble, F-38700 France

2 - Department of Pathology, Brigham and Women's Hospital, Boston, M A, United States.

3 - EMBL Grenoble, BP 181, 6 rue Jules Horowitz, 38042 Grenoble Cedex 9, France

4 - Department of Pharmacology and Molecular Sciences; Department of Oncology ; Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

5 – Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan

*Corresponding author:

khochbin@ujf-grenoble.fr

Tel : (33) 4 76 54 95 83 ; Fax : (33) 4 76 54 95 95

Supplementary Material and methods

Gene reporter assays and FACS analysis

Cos7 cells were co-transfected for 36 hours with β-Galactosidase reporter plasmid and *p21*-promoter luciferase reporter plasmid either alone or with p53-Flag and increasing amounts of BRD4-NUT-HA plasmid. The total amount of transfected plasmid was equilibrated with pcDNA-HA plasmid between the different conditions. Cells were then washed and incubated 15 min at RT with Luciferase lysis buffer (1% Triton X-100, 10% Glycerol, 2mM EDTA, 25mM Tris-HCl pH 7.8). Luciferase activity was measured directly after mixing 10µL of cellular extract with 50µl of Luciferase substrate (Luciferase assay kit, Stratagene). Normalization was made by measuring the β-Gal activity after 1hour incubation of 7.5µL of cellular extract with 50µL of β-Gal substrate (β-Gal detection kit II, Clontech).

For apoptosis FACS analysis, HCC2429 cells were treated with scrambled or BRD4-NUT siRNA for 24h, with or without 20µM etoposide for 16h. Cells were then collected, permeabilized and fixed for one hour on ice in methanol, washed and incubated with PE-A conjugated active Caspase 3 antibody (PE-active caspase 3 apoptosis kit, BD Biosciences). Cells were then analyzed using a BD Biosciences LSRII cytometer.

Immunoprecipitation, HAT assay, GST pull down, insect cell culture, baculovirus cultivation and protein purifications

Cos7 cells were transfected and, 36 hours post-transfection, a cell extract was prepared in LSDB 200 or 500mM KCL (Glycerol 20%, MgCl2 3mM, Hepes pH7.9 50mM, KCL 200 or 500mM, DTT 0,5mM, PMSF 0,5mM, NP40 0,1%, Complete EDTA-free 1x, TSA 0,05µg/mL). The cell extract was incubated with anti-HA rat (Roche) for 2 hours at 4°C, then incubated with protein G sepharose beads for 2 hours at 4°C with rotation. After intensive

wash with LSDB buffer, beads were used either for HAT assays or directly eluted with Laemmli buffer for immunoblotting.

NUT-GST fragments were produced in IPTG-induced BL21 transformed with Pet28 plasmids. Bacteria lysates were incubated over night at 4°C with Glutathione sepharose beads (GE Healthcare). Beads were intensively washed and then incubated with extracts from Cos cells expressing proteins of interest or with purified p300 proteins. After intensive washes, the beads were used for HAT assays or directly eluted with Laemmli buffer.

For HAT assays, after immunoprecipitation, beads were incubated 2h at 30°C with 25µL of HAT buffer (Glycerol 10%, Tris HCL pH8 25mM, NaCl 100mM, EDTA 0,1mM, DTT 1mM, PMSF 0,2mM, Complete EDTA-free 1x, TSA 0,05µg/mL, Acetyl Coenzyme A ^{14}C 1µL at 20nCurie/µL (MPbio)) and 1µg purified free histones or purified histone H3. After electrophoresis, coomassie coloration and incubation in "Amplify" solution (Amersham), the gels were dried and radioactivity was detected using a PhosphoImager (Molecular Dynamics).

Purified NUT F1c was obtained by eluting Glutathione sepharose beads-bound GST fusion proteins with Glutathione 25mM, 50mM Tris pH8 (Sigma). Glutathione was then removed and NUT F1C concentrated by using Amicon Ultra 10,000 NMWL (Millipore).

Purified histone H3 was obtained by a production in IPTG-induced BL21 and by histone extraction (acid extraction protocol from Abcam). p300L (324-2094) was cloned in the vector pFASTBAC1 (Invitrogen). The recombinant baculoviruses were propagated in Sf21 insect cells with SF900-II SFM insect cell medium (Invitrogen). For production of recombinant p300L, Hi5 insect cells were infected with recombinant baculovirus at a multiplicity of infection of 5.0, maintained in shake flasks at 28 °C, and harvested by centrifugation at 72 h postinfection.

For purification of p300L, -recombinant baculovirus-infected cell pellets were lysed by freezing and then thawed in equilibration buffer containing 20 mM Tris (pH 8.0), 300 mM

NaCl and Complete Protease Inhibitors EDTA-free (Roche Applied Science). The cells were homogenized by sonication and the lysate clarified by centrifugation at 180,000 x g for 40 min. The clarified insect cell lysate was incubated with pre-equilibrated Anti-FLAG M2-agarose affinity gel (Sigma) for 1 h and then applied to a 50-ml Econo column (Bio-Rad). The resin was washed with 200 ml of equilibration buffer, and then the bound protein was eluted using 0.1 mg/ml triple FLAG peptide (sequence, MDYKDHDADYKDHDIDYKDDDDK) dissolved in equilibration buffer.

Fractions were concentrated using an Amicon-ultra 30 unit (Amicon, Inc.) and further purified on a Hi-Load 16/60 Superdex 200 gel filtration column (GE Healthcare) equilibrated in 20 mM HEPES (pH 7.5), 150 mM NaCl, 1 mM dithiothreitol (DTT). The final protein was concentrated to 10 mg/ml; aliquots were flash-frozen in liquid N2 and stored at -80 °C. For p300S (aa 1045-1666), the construct was cloned in pET Duet with a STREP tag and expressed in *E.coli* BL21 (DE3) using autoinducing media ZYP-5052 (Studier, 2005). Cell pellets were resuspended in lysis buffer (20mM Tris, pH 8.0, 300 mM NaCl, 1mM DTT, 25 µM $ZnCl_2$ and Complete Protease Inhibitors EDTA-free (Roche Applied Science) and lysed by sonication. The protein was purified from crude extract by STREP-affinity chromatography followed by Gel Filtration on a Hi-Load 16/60 Superdex 200 gel filtration column (GE Healthcare) equilibrated in 20 mM HEPES (pH 7.5), 300 mM NaCl, 25 µM $ZnCl_2$, 1 mM dithiothreitol (DTT). The final protein was concentrated to 1 mg/ml.

Supplementary Figure legends

Supplementary Figure S1: Colocalization of hyperacetylated histone H4 with the BRD4-NUT foci. Cos7 cells were transfected and processed as described in Figure 1, but a higher magnification of the BRD4-NUT foci shown in Figure 1A is presented. This Figure also shows the co-detection of the longer isoform of BRD4 (flBRD4) either with H4ac or NUT (B and C, indicated). Bar, 10μm.

Supplementary Figure S2: Association of hyperacetylated but transcriptionally inactive chromatin with BRD4-NUT foci. (A) and (C) Cos7 cells were transfected with a GFP-BRD4-NUT expressing vector and the presence of the indicated histone marks in the foci was monitored using a series of specific antibodies (indicated). (B) In an experiment similar to (A), the presence of RNA polymerase II (RNA pol II) and its phosphorylated forms on serine 2 (CTD S2Ph) and serine 5 (CTD S5Ph) in BRD4-NUT foci was analysed using the indicated specific antibodies in Cos7 cells transfected with GFP-BRD4-NUT. (D) Endogenous BRD4-NUT foci were co-detected with H3K14ac and H3K27ac marks in HCC2429 cells. Bar, 10μm.

Supplementary Figure S3: Specific colocalization of BRD4-NUT foci with p300. Cos7 cells were transfected with the GFP-BRD4-NUT expression vector and the indicated endogenous HATs were visualized by indirect immunofluorescence using the corresponding antibodies. Bar, 10μm.

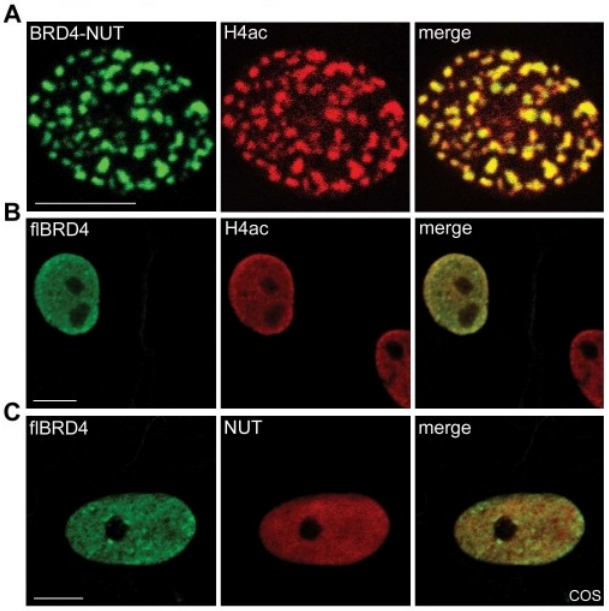
Supplementary Figure S4: Specific activation of p300 catalytic activity by NUT. The HAT activity of purified p300 was monitored as described in the legend of Figure 4 using the indicated amounts of NUT fragments (F1c, left panel and F1b, right panel).

Supplementary Figure S5: Disruption of BRD4-NUT foci after the inhibition of class I/II histone-deacetylases. (A) HCC2429 cells were treated with 50ng/mL TSA for the indicated times and BRD4-NUT and p300 were visualized with their respective antibodies. (B) In a similar experiment as in "(A)", HCC2429 were treated for the indicated times and BRD4-NUT, p21 and ACTIN were visualized. (C) The amount of *GADD45* and *BRD4-NUT* mRNA was monitored in non-treated HCC2429 cells or cells treated for 5 hours with 50ng/mL TSA, normalized by comparison with *GAPDH* mRNA, and represented as the fold variation with respect to non treated cells. Data represent the mean +/- standard deviation of at least two independent experiments performed in triplicates. (D) HCC2429 cells were treated for 5 hours with the indicated molecules and BRD4-NUT, p53, p21 and ACTIN visualized as above.

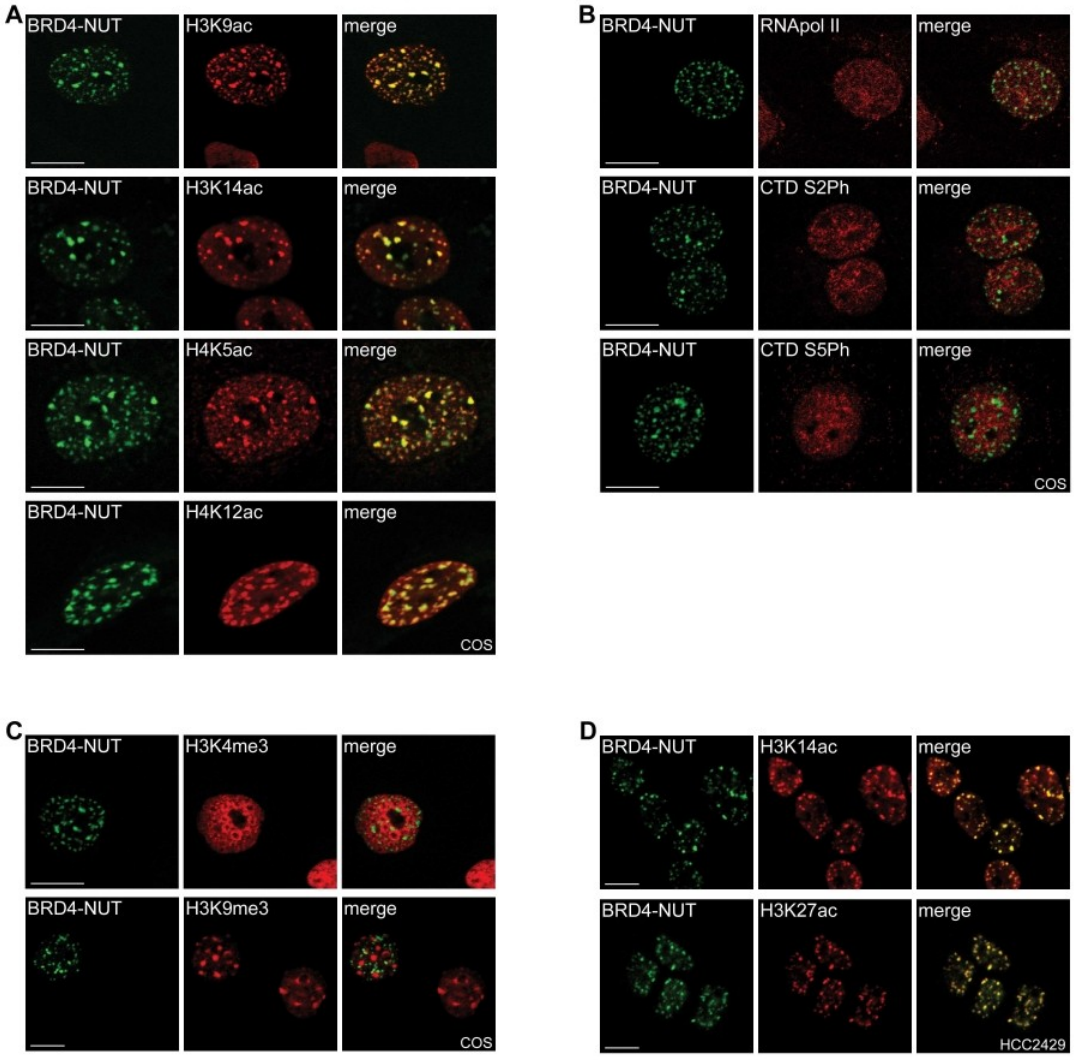
Supplementary Figure S6: p53 inactivation in BRD4-NUT expressing cells. (A) Total p53 (p53) or K373 and K382 acetylated p53 (p53Kac) were detected in HCC2429 cells. HCC2429 cells were treated either with scrambled and two anti-BRD4-NUT siRNAs and acetylated p53 and total p53 were detected as indicated. Control panels show non siRNA-treated cells. Bar, 10μm. (B) HCC2429 cells were treated either with scrambled and two anti-BRD4-NUT siRNAs in presence or absence of etoposide (20μM for 15 hours) and the apoptotic cells were visualized by immunofluorescent detection of the active caspase 3 and FACS analysis. The percentage of cells containing active caspase 3 with respect to the total number of analyzed cells is indicated (% apoptotic cells). The histogram indicates the mean values of three independent experiments, and two representative FACS analyses are shown (right panel).

Supplementary Figure S7: Inhibition of p53 activity by BRD4-NUT. (A) Cos7 cells were transfected with GFP-BRD4-NUT and p53, acetylated p53 and BRD4-NUT were visualized with the respective antibodies as indicated. (B) The same cells were transfected with a *p21* promoter-luciferase reporter construct, p53-Flag plasmid and increasing amounts of BRD4-NUT-HA plasmid, as well as a CMV-βGal construct for normalization, as indicated. 36 hours post-transfection, the luciferase activity was measured in independent triplicates, normalized with respect to β-Galactosidase activity and represented as fold variation compared to Luciferase activity in the absence of p53.

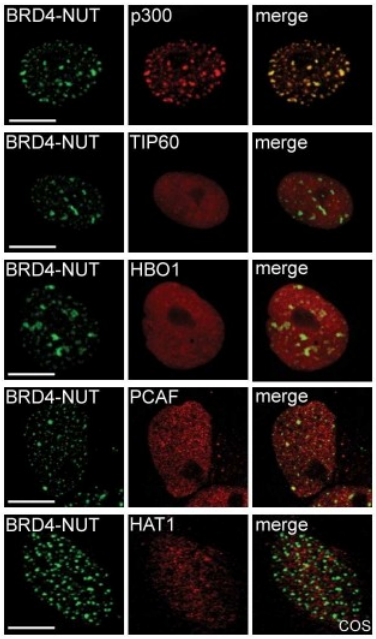
Suppl. Figure 1 Reynoird *et al.*



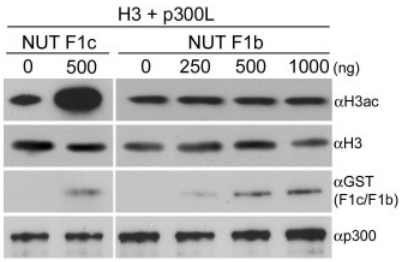
Supplementary Figure 2 Reynoird *et al.*



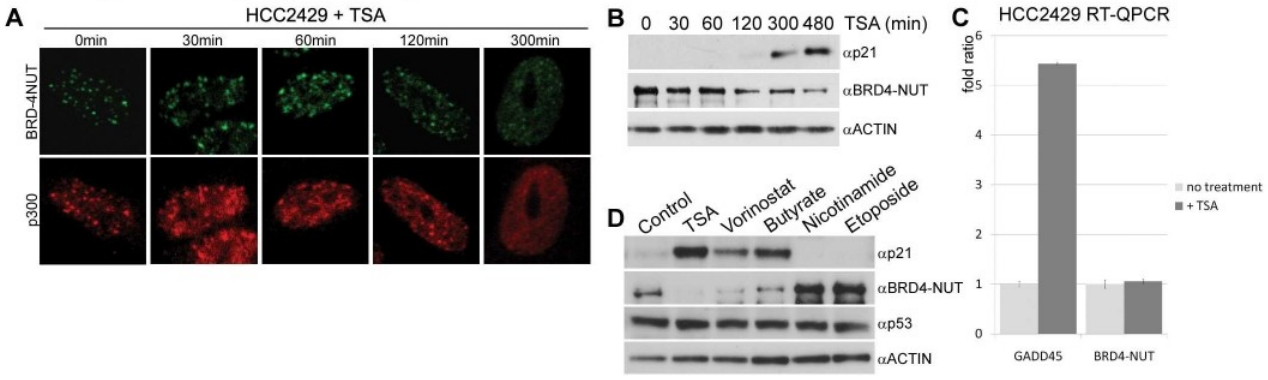
Supplementary figure 3 Reynoird *et al.*



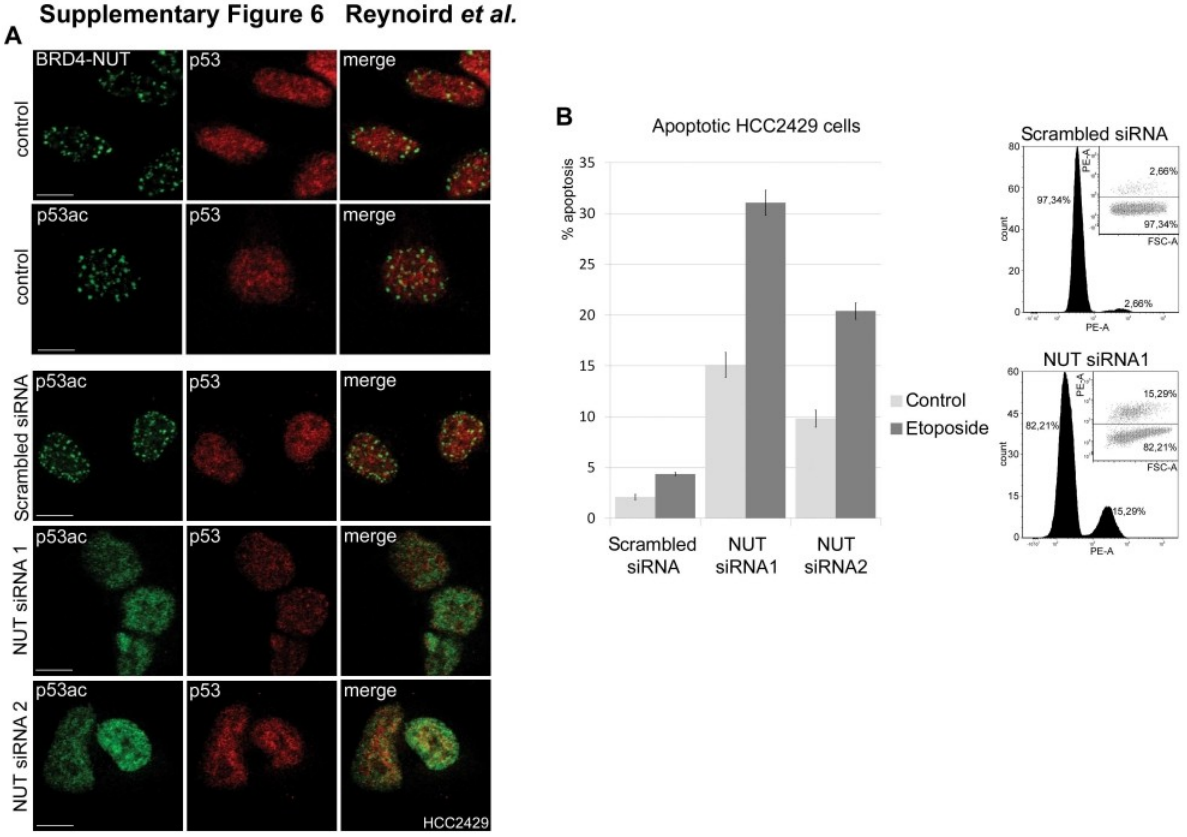
Supplementary Figure 4 Reynoird *et al.*



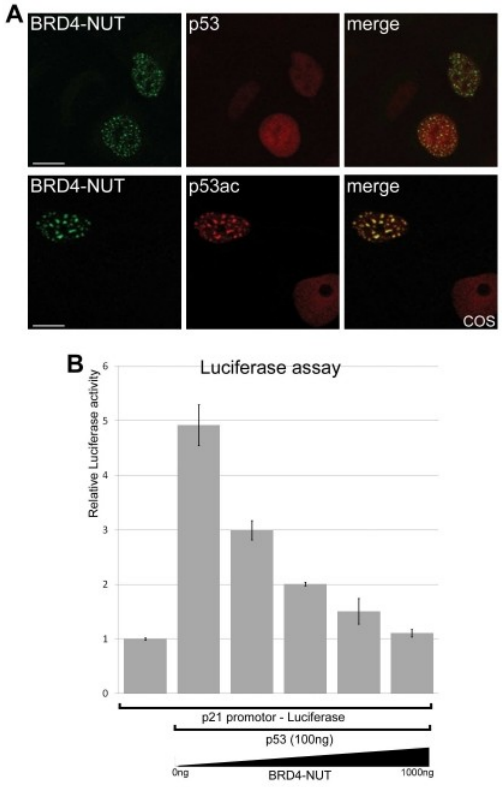
Supplementary Figure 5 Reynoird *et al.*



Supplementary Figure 6 Reynoird *et al.*



Supplementary Figure 7 Reynoird *et al.*



AKNOWLEDGEMENTS

Panne group:	Daniel Panne
	Carmen Aguilar, Joanna Andrecka, Juliette Devos, Amédé Larabi, Emily Newman
Collaborators:	Saadi Khochbin, Nycolas Reynoird (IAB, Grenoble)
TAC commettee:	Stephen Cusack, Saadi Khochbin, Andreas Ladurner, Daniel Panne
For the precious help:	Max Nanao, Andres Palencia, Danielle Desravines, Federico Forneris
ESRF:	ID14-4, ID14-3, ID23-1 beamlines
SOLEIL synchrotron:	SWING beamline
EMBL facilities:	MS, DNA sequencing, HTX lab
IBS facilities	MS, N-ter sequencing, 1D-NMR

All EMBL people, past and present

My family, my love and my friends, who gave me the strength not to give up: my success is all yours.

“Keep the vibe”

“Garde la peche”