Abstract: The acetylation/acylation (ac(et)ylation) of lysine side chains is a dynamic post-translational modification (PTM) regulating fundamental cellular processes with implications on the organisms’ ageing process: metabolism, transcription, translation, cell proliferation, regulation of the cytoskeleton and DNA damage repair. First identified to occur on histones, later studies revealed the presence of lysine ac(et)ylation in organisms of all kingdoms of life, in proteins covering all essential cellular processes. A remarkable finding showed that the NAD⁺-dependent sirtuin deacetylase Sir2 has an impact on replicative lifespan in Saccharomyces cerevisiae suggesting that lysine acetylation has a direct role in the ageing process. Later studies identified sirtuins as mediators for beneficial effects of caloric/dietary restriction on the organisms’ health- or lifespan. However, the molecular mechanisms underlying these effects are only incompletely understood. Progress in mass-spectrometry, structural biology, synthetic and semi-synthetic biology deepened our understanding of this PTM. This review summarizes recent developments in the research field. It shows how lysine ac(et)ylation regulates protein function, how it is regulated enzymatically and non-enzymatically, how a dysfunction in this post-translational machinery contributes to disease development. A focus is set on sirtuins and lysine acyltransferases as these are direct sensors and mediators of the cellular metabolic state. Finally, this review highlights technological advances to study lysine ac(et)ylation.

Keywords: ageing; KDAC; longevity; lysine acetylation; lysine acyltransferases; lysine deacetylases; sirtuins; synthetic biology.

Introduction

The human genome encodes approximately 20,000–25,000 proteins (International Human Genome Sequencing 2004). The diversity of this proteome can be substantially enlarged by processes such as alternative splicing and post-translational modifications (PTMs). PTMs range from attachment of chemical groups such as observed in phosphorylation of serine, threonine or tyrosine side chains dynamically regulated by kinases and phosphatases, to attachment of small proteins such as ubiquitin for regulation of protein turnover, subcellular localization and signal transduction. Moreover, proteolytic processing of proteins is a well-known PTM as exemplified by the activation of coagulation factors or the activation of digestive enzymes from inactive zymogens. Proteins can be lipidated to regulate membrane binding/recruitment by forming irreversible thioethers between cysteine side chains and prenyl-groups or by forming reversible thioesters by attachment of fatty acids (palmitoylation, myristoylation) to cysteine side chains (Chen et al. 2018). Apart from reversible lipidation on cysteine side chains, lysine side chains can undergo reversible lipidation/acetylation.
formation of an isopeptide bond between the ε-amino group of the lysine side chain and the ε-carboxyl group of the terminal diGly motif of Ub/Ubl (Celen and Sahin 2020; Swatek and Komander 2016). The ubiquitin can be removed by deubiquitinases (DUBs). Moreover, lysine side chains can be modified by mono-/di-/tri-methylation, which alters the side chain sterically and enables the binding of proteins containing specific methyl-lysine reader domains. Tudor domains enable binding to di- and trimethylated lysines, while MBT domains bind to mono-/di-methylated lysines and PHD domains bind to various methylation states (mono-/di-/tri-methyl-lysine) (Wang et al. 2020a,b). PHD domains are zinc containing domains that often occur as tandem PHD domains. Chromodomains, PWWP- and WD domains bind to methylated lysine residues. All of these domains target various methylated lysines in histone tails and are involved in chromatin remodeling and regulation of gene expression. Lysine side chains can be modified by acetylation and various other acylations such as the charged succinylation, malonylation or glutarylation and aliphatic acylations such as formylation, propionylation, butyrylation, myristoylation or palmitoylation (Figure 1(A)). Recently, lysine 2-hydroxyisobutyryltransferylation occurring on histones with high stoichiometry inducing structural changes in the chromatin was shown to regulate glycolysis and being catalyzed by the KAT p300 (Figure 1(C)) (Dai et al. 2014; Huang et al. 2018a,b, 2020). In contrast to mono-/di-/tri-methylation, the acetylation/acylation of lysine side chains results in neutralization of the positive charge occurring at the lysine side chain at physiological pH.

Lysine acetylation was identified together with methylation by Phillips in 1963 to occur on histones and in 1964 Allfrey reported that acetylation and methylation of histones regulates RNA synthesis, suggesting that it has an impact on gene expression (Allfrey and Mirsky 1964; Allfrey et al. 1964; Phillips 1963). Today it is known that not only histones are lysine-acetylated, but in fact the number of lysine acetylation sites rivals that of phosphorylation sites and thousands of proteins are acetylated in all cellular compartments and it is present in all kingdoms of life (Choudhary and Mann 2020; Choudhary et al. 2009, 2014; Hansen et al. 2019; Kouzarides 2000; Lundby et al. 2012; Weinert et al. 2011, 2015, 2017, 2018).

This review highlights the physiological role of lysine acetylation. We discuss the regulation of lysine acetylation mechanistically with respect to cellular metabolic pathways. It summarizes how lysine acetylation is regulated in cells by writers, the lysine acetyltransferases (KATs), erasers, the lysine deacetylases (KDACs), and how it is targeted by reader domains such as bromodomains (BRDs) and YEATS domains (Figure 1(A)). We will discuss how lysine acetylation is involved in disease development and will put an emphasis on recent technological advances to study lysine acetylation.

Enzymatic regulation of lysine acetylation by classical KDACs, sirtuins and KATS

Classical Zn²⁺-dependent lysine deacetylases

After the discovery of lysine acetylation on histones in 1963, enzymatic activity to remove acetyl-groups was discovered in calf thymus in 1969 (Inoue and Fujimoto 1969; Phillips 1963). However, the isolation of the first lysine deacetylase HDAC1 took until 1996 (Taunton et al. 1996). Later studies showed the presence of two lysine deacetylase families, which are structurally and based on their primary sequences unrelated: 1.) classical histone/lysine deacetylases (HDACs/KDACs; Zn²⁺-dependent catalytic mechanism) and sirtuin deacetylases (sirtuins; SIRT: silent information regulator; NAD⁺-dependent catalytic mechanism) (Seto and Yoshida 2014; Yoshida et al. 2017). As the deacetylases were first characterized as histone deacetylases, the name was abbreviated with HDAC. However, as recent data show that these enzymes have many substrates besides histones they are also called lysine deacetylases (KDACs) (Kumar et al. 2021; Peng and Seto 2011; Sadoul et al. 2011). Today it is known that 11 classical deacetylases are encoded in the human genome (Figure 2(A)) classified into three classes: class I (with HDAC1/2/3/8 showing sequence similarity to yeast Rpd3), class II (subclass IIa with HDAC4/5/7/9 and subclass IIb with HDAC6/10 showing sequence similarity to yeast Hda1) and class IV (with HDAC11 showing sequence similarity to class I and class II enzymes) (Figure 2(A)) (de Ruijter et al. 2003; Haberland et al. 2009). These enzymes of classes I, II and IV are structurally similar and use a catalytic Zn²⁺ ion to hydrolyze the amide bond to remove acetyl-groups from the ε-amino group of lysine side chains (Figure 2(B) and (C)) (Lombardi et al. 2011).

The catalytic mechanism involves the coordination and polarization of a catalytic water molecule by the catalytic Zn²⁺-ion (Figure 2(C)). This Zn²⁺-ion together with a Tyr residue (HDAC8: Y306), which is replaced by a histidine in class IIa KDACs, also interact with the carbonyl oxygen of the acetyl group to polarize the carbonyl increasing the electrophilicity at the carbonyl carbon atom.
for nucleophilic attack of the catalytic water molecule (Park and Kim 2020). Mutational studies suggest that two His residues of a tandem His-motif are important for catalytic activity. One His residue (HDAC8: H143), polarized and oriented by an Asp (HDAC8: D183), acts as general base, abstracting a proton from the water molecule. A second His (HDAC8: H142) again polarized and oriented by another Asp (HDAC8: D176) acts as electrostatic catalyst. This His orients and polarizes the catalytic water molecule, which attacks by a nucleophilic reaction the electrophilic carbonyl carbon of the acetyl-group. A tetrahedral intermediate is formed, which is stabilized by the Tyr residue (HDAC8: Y306). This intermediate resolves by proton transfer from a His (HDAC8: H143), which acts as catalytic acid to the deacetylated lysine side chain (Seto and Yoshida 2014; Somoza et al. 2004). This finally results in release of an acetate and the deacetylated lysine (Figure 2(C)).

In HDAC enzymes of class IIa, the catalytic Tyr residue (HDAC8: Y306) is replaced by a histidine. This is most likely one reason why class IIa enzymes show only low deacetylase activities, which can be increased by replacing His by Tyr (Lahm et al. 2007). It is postulated that this low basal activity in HDACs of subclass IIa has evolved during evolution to ensure a high specificity towards a narrow range of substrates. Along this line, it was found, that subclass IIa enzymes are highly active in deacetylating trifluoroacetylated peptides (Lahm et al. 2007).

Compared to kinases, a quite low number of classical deacetylases are encoded in the human genome. This opens the question of how substrate specificity is created (Milazzo et al. 2020; Sugiyama et al. 2019). Recent data shows that the classical HDACs HDAC1/2/3 of class I are often part of nuclear multi-protein complexes, which regulate their activity and substrate recognition (Codina et al. 2005; Kalin et al. 2018; Kelly et al. 2018; Millard et al. 2020; Turnbull et al. 2020; Wang et al. 2020a,b). These complexes mainly regulate gene expression by targeting different sets of acetylated histones as substrates. Several of these complexes were structurally and functionally characterized: NuRD complex: HDAC1/2; CoREST: HDAC1/3; Sin3: HDAC1/2, MIDAC: HDAC1/2, NCoR/SMRT: HDAC3. The activities of the HDACs in these complexes can be modified by different co-factors. The formation of these complexes, in which several deacetylase activities are present, impedes therapeutic targeting (Kalin et al. 2018; Mathias et al. 2015; Segre and Chiocca 2011; Yang and Seto 2007).

Classical deacetylases are either located in the nucleus or in the cytosol (Figure 2(A)) and are regulated by post-translational modifications such as phosphorylation, ubiquitination, SUMOylation, glycosylation and acetylation (Gallinari et al. 2007; Mathias et al. 2015; Segre and Chiocca 2011). HDACs/KDACs control fundamental cellular processes such as gene expression, transcription, translation,
autophagy, chromatin remodeling, cell cycle and cytoskeleton dynamics (Gallinari et al. 2007; Lam et al. 2020; Liao et al. 2020; Mrakovcic et al. 2017). They have histones and many non-histone proteins as substrates and they are reported to contribute to the development of severe diseases such as cancer, neurodegenerative diseases, metabolic and inflammatory defects, cardiovascular diseases and they affect the overall organisms’ ageing process (Janczura et al. 2018; Li et al. 2016a,b; McIntyre et al. 2019; Pasyukova and Vaiserman 2017; Shukla and Tekwani 2020; Yoon and Eom 2016). The expression of HDAC genes is affected in several tumor types (Glozak and Seto 2007; Lucio-Eterovic et al. 2008; Wang et al. 2020a,b), and in turn, altered HDAC protein levels have consequences on the expression of tumor suppressor genes or oncogenes by affecting histone-acetylation states (Feng et al. 2007, 2013; Glozak and Seto 2007; Lucio-Eterovic et al. 2008; Wang et al. 2020a,b).
neurotoxicity in a Parkinson disease model. Tubastatin A protected against α-synuclein dependent neurotoxicity in a Parkinson’s disease model (Supplementary Material Figure 1(A)) (Francelle et al. 2020; Shukla and Tekwani 2020). To this end, classical HDACs are potent targets for therapeutic interventions. As an example, HDAC inhibitors were shown to counteract DNA repair in rapidly proliferating cells (Nikolova et al. 2017). Several inhibitors for classical HDACs were developed and some are approved by the FDA (Food and Drug Administration) (Beyer et al. 2019; Nebbioso et al. 2017). These can be classified by their chemical structures: hydroxamic acids, benzamides, short fatty acids, cyclic peptides (Supplementary Material Figure 1(A)) (Kim and Bae 2011; Li and Seto 2016; Sanaei and Kavoosi 2019; Xu et al. 2007). The hydroxamic acid inhibitors trichostatin A (TSA) and SAHA (vorinostat) show a nanomolar inhibitory potency and act by chelating the catalytic Zn$^{2+}$ ion. While TSA was a natural isolate from Streptomyces, SAHA had improved pharmacokinetic properties compared to TSA and was the first KDAC inhibitor that was used as potent anti-cancer drug (Marks 2007; Xu et al. 2007). Another important class of HDAC/KDAC inhibitors are benzamides such as MS-275 (entinostat), which also chelate the Zn$^{2+}$ ion (Supplementary Material Figure 1(A)) (Simonini et al. 2006). Most of the inhibitors were either pan-inhibitors affecting the activities of almost all classical deacetylases or they target a subclass. Several attempts were initiated to develop selective and potent inhibitors that target individual HDAC isoforms. Selective and potent inhibitors were obtained for HDAC6, i.e. the hydroxamate tubacin (IC$_{50}$: 4 nM), and HDAC8, i.e. the hydroxamate PCI-34051 (IC$_{50}$: 10 nM) (Supplementary Material Figure 1(A)) (Bieliauskas and Pflum 2008; Estiu et al. 2010; Itoh et al. 2008; Neiband et al. 2020; Zhang and Xu 2015). Other specific inhibitors were recently developed for HDAC6 (Leonhardt et al. 2018; Noack and Kramer 2017). The highly similar active sites as well as the similar catalytic mechanisms of the different HDACs hamper the development of specific inhibitors for individual HDAC isoforms.

**NAD$^{+}$-dependent sirtuin deacetylases**

Sirtuins, first identified in Saccharomyces cerevisiae in 1985, were shown to be involved in silencing of transcription of mating-type loci HML and HMR, which contain the mating-type information. Furthermore, it was discovered that sirtuins are involved in protection of telomers (Ivy et al. 1985; Rine and Herskowitz 1987). Later, in 1999, it was shown, that the gene locus encompassing sir2/3/4 in yeast is involved in regulation of replicative lifespan in yeast (Kaeberlein et al. 1999). In fact, sir2 and the corresponding gene product SIR2 alone is sufficient to significantly increase longevity while its genomic deletion results in a decrease of the replicative lifespan in yeast (Kaeberlein et al. 1999). This was a remarkable finding as this showed, that a single gene product is sufficient to regulate a highly complex process such as organisms’ ageing. Later it was shown that SIR2 (silent mating type information regulator 2) constitutes an NAD$^{+}$-dependent histone deacetylase (Imai et al. 2000). Already in the 1930s, experiments with mice and rats revealed that caloric restriction, without malnutrition, has a beneficial impact on lifespan (McCay et al. 1975, 1989). However, more recent evidence clarified it is the composition of the diet rather than its calorie content that mediates the longevity effect (Piper et al. 2011). Sirtuins are believed to act as mediators of this prolongevity effect of caloric/dietary restriction at the molecular level (Lee et al. 2019). The beneficial impact of sirtuins on lifespan is not undisputed and might depend on the level of the tissue of sirtuin expression (Fabricio et al. 2005; Lee et al. 2019). However, data obtained with all model organisms, S. cerevisiae, Caenorhabditis elegans, Drosophila melanogaster and Mus musculus, shows that sirtuins either directly affect the organisms’ ageing process or health span or they are mediators of caloric/dietary restriction at the molecular level (Grabowska et al. 2017; Guarente 2013; Zullo et al. 2018). As sirtuins use NAD$^{+}$ as a stoichiometric co-substrate for deacetylation, they act as sensors of the cellular NAD$^{+}$ level (Figure 3(B), Supplementary Material Figure 1(E)). As such they are able to translate the metabolic state into alteration of the acetylation state and in turn of the activities of sirtuin substrate proteins. Following the identification of sirtuins in yeast, sirtuins were identified in bacteria and in mammals (Figure 3(A)) (Michan and Sinclair 2007). Most Gram-positive and Gram-negative bacterial species encode for a single sirtuin deacetylase (Bacillus subtilis: SrtN; Escherichia coli: CobB) (Gardner and Escalante-Semerena 2009; Tsang and Escalante-Semerena 1998; Zhao et al. 2004). As CobB is the only sirtuin deacetylase in E. coli, it most likely developed during evolution as a highly...
promiscuous enzyme protecting against a decrease in protein functionalities due to increase in chemical ac(et)ylation caused by elevated levels of acetyl-CoA or acetyl-phosphate as a consequence of processes such as metabolic fuel switching (Figure 4(A) and (D)). Mammals encode for seven sirtuins, which differ in their subcellular localization, their activities and the sequences surrounding the conserved catalytic core in the N- and C-termini (Figure 3(A)). SIRT1 corresponds to the yeast Sir2. SIRT1, SIRT6 and SIRT7 are predominantly nuclear, SIRT7 is located in the nucleoli, SIRT2 is predominantly cytosolic and SIRT3/4/5 are located within the mitochondrial matrix (Figure 3(A)). Later studies suggested that SIRT1 does have some cytosolic substrates and it shuttles between nucleus and cytosol (Bai and Zhang 2016; Byles et al. 2010; Sun and Fang 2016; Tanno et al. 2007). SIRT1 has two nuclear localization signals (NLSs) and two nuclear export signals (NESs) and it was supposed that upon mutation of these sequences SIRT1 loses its shuttling capacity (Tanno et al. 2007). Of the seven sirtuins, only SIRT1/2/3 are robust deacetylases, i.e. its primary activity is to deacetylate lysine side chains. However, they are capable, albeit with lower activity, of removing longer fatty acid derived acyl groups from lysine side chains in vitro (Figure 3(A)) (Feldman et al. 2015). The other sirtuins show strongly reduced deacetylation activities and have preferences for other acylations occurring at lysine side chains (Bheda et al. 2016; Denu and Gottesfeld 2012; Dittenhafer-Reed et al. 2011; Feldman et al. 2012, 2015).

Our data on natively-folded and site-specifically lysine-acetylated substrate proteins suggest that sirtuins can show a remarkable level of substrate specificity, which is created by both, the primary sequence and the three-dimensional structure (Knyphausen et al. 2016a,b). In contrast to earlier reports, our data shows that natively-folded and site-specifically lysine-acetylated sirtuin substrate proteins are efficiently deacetylated by sirtuins if the acetylation sites are located in loop regions rather than in elements of pronounced secondary structure such as α-helices and β-strands (Choudhary et al. 2009; Knyphausen et al. 2016a,b). This is in agreement with post-translational phosphorylation, which is of functional importance occurring in loops accessible for kinases and phosphatases (Knyphausen et al. 2016a,b).

The robust deacetylases SIRT1 and SIRT2

The nuclear deacetylase SIRT1

SIRT1 regulates a broad range of important transcription factors and coregulators, including p53, FOXO, PGC-1α and c-Myc (Kobayashi et al. 2005; Mao et al. 2011; Ren et al. 2019; Rodgers et al. 2005; Solomon et al. 2006). SIRT1 is involved in fundamental cellular processes such as integrity of chromatin, metabolism, autophagy, mitophagy, apoptosis, cell cycle and DNA-damage response (Chang and Guarente 2014). By modulating the acetylation status of histones, SIRT1 regulates chromatin structure and function to modulate transcriptional activity (Bosch-Preseguere and Vaquero 2014; Imai et al. 2000; Vaquero et al. 2004, 2007). Moreover, SIRT1 deacetylates the protein light chain 3 (LC-3) and the kinase PINK1, thereby modulating autophagy and mitophagy, respectively (Fang et al. 2014; Huang et al. 2015; Kitagishi et al. 2017; Ou et al. 2014; Yao et al. 2018). SIRT1 is the most widely studied sirtuin and there are many excellent reviews that set a focus specifically on SIRT1 and its role in cellular physiology, health and disease (Bonkowski and Sinclair 2016; Chalkiadaki and Guarente 2015; Chang and Guarente 2014; Guarente 2013; Haigis and Sinclair 2010; Imai and Guarente 2016; Rizzi and Roriz-Cruz 2018).

The cytosolic deacetylase SIRT2 can act as de-fatty acylase

For the cytosolic SIRT2 it was shown that it has activity to remove long chain fatty acids, i.e.effectively removing lysine myristoylation, acting as de-fatty acylase, next to its robust deacetylase activity (Feldman et al. 2015; Hong et al. 2020; Jing et al. 2017; Kosciuc and Lin 2020; Teng et al. 2015; Wang et al. 2017a,b; Zhang et al. 2019a,b). SIRT2 is involved in regulation of the cell cycle, in the regulation of the cytoskeleton, in inflammatory processes and in cellular response to reactive oxygen species (ROS) (Inoue et al. 2007a,b; Lemos et al. 2017; Min et al. 2018). It deac(et)ylates, besides others, α-tubulin and the RhoA regulator RhoGDIα, HIF-1α and other non-histone substrates, such as K-Ras-4b, RalB, Ran, FOXO3, NF-κB and lactate-dehydrogenase (de Boor et al. 2015; Inoue et al. 2007a,b; Jing et al. 2017; Kuhlmann et al. 2016a,b; Kuhlmann et al. 2016a,b; Seo et al. 2015; Spiegelman et al. 2019). Additionally, it was observed that SIRT2 affects the acetylation status in mitochondria and it dysregulates mitophagy (Liu et al. 2017). Recently, it was reported that SIRT2 can actively be imported into the nucleus dependent on its C-terminus and on nuclear import receptors (Eldridge et al. 2020). It was shown earlier that SIRT2 shuttles between cytosol and nucleus and is localized in the nucleus during mitosis or upon bacterial infections (North and Verdin 2007a,b; Vaquero et al. 2006). Deletion of sirt2 impairs insulin resistance in mice fed with a high fat diet suggesting that it results in an impaired insulin action exerting a direct role
in regulation of metabolism (Lantier et al. 2018). Notably, this effect was not connected to alterations of the cytosolic acetylation level and pattern but it alters the mitochondrial acetylation state (Lantier et al. 2018; Liu et al. 2017). These results show that SIRT2 has a direct role on the regulation of mitochondrial physiology. One recent study showed that SIRT2 is localized to the inner mitochondrial membrane of cells in the central nervous system in mouse brains (Liu et al., 2017). It still needs to be elucidated how SIRT2 mechanistically localizes to the inner mitochondrial membrane as it lacks an obvious mitochondrial targeting sequence.

Figure 3: Sirtuins are NAD$^+$-dependent lysine-deacetylases. (A) Mammals encode seven sirtuins with different subcellular localization. SIRT1 is predominantly nuclear, some reports state that it shuttles between cytosol and nucleus under specific conditions. SIRT2 is cytosolic, SIRT3/4/5 contain an N-terminal mitochondrial targeting sequence and are mitochondrial, SIRT6 is nuclear and SIRT7 is located in the nucleolus. For SIRT5 a cytosolic role was suggested. Phylogenetic studies suggest the mammalian sirtuins to be classified into four classes, sometimes with subclasses, as indicated. A fifth class is formed by sirtuins from Gram-positive bacteria or archaea. (B) Sirtuins use NAD$^+$ as stoichiometric co-substrate for catalysis. The carbonyl-oxygen of the acetyl-lysine attacks the C-1′ of the ribose of NAD$^+$ as a nucleophile resulting in release of nicotinamide as a leaving group and formation of a C-1′-O-alkylamidate intermediate. This intermediate is resolved in several steps, including a hydrolysis reaction, as indicated, finally resulting in formation of the deacetylilated lysine and 2′-O-acetyl-ADP-ribose, which can be converted to yield 3′-O-acetyl-ADP-ribose. (Figure redrawn and modified from Ali et al. 2018; Smith and Denu 2006a,b; Teixeira et al. 2020). (C) Structure of the catalytic domain of human SIRT3 in complex with carba-NAD$^+$ and an acetylated acetyl-CoA-synthetase 2 peptide. Shown are the domain boundaries and structural features in different colors as indicated. The NAD$^+$ is bound by a Rossmann-fold domain consisting of a six-stranded parallel β-sheet flanked by several α-helices, the exact number dependent on the sirtuin. The phosphates of NAD$^+$ are bound by a conserved Gly-X-Gly motif. The sirtuins belong to the inverted Rossmann-fold enzymes with the adenine-moiety binding in the A-site located at the C-terminal half of the fold and the nicotinamide moiety in the N-site located in the N-terminal half. The ribose is bound by the B-site. The active site is located between the Rossmann-fold and the Zn$^{2+}$-binding/small helical domain (PDB code: 4FVT). (The figure was generated with PyMOL v.2.3.4 Schrödinger, LLC, New York, NY, USA).
Figure 4: Pathways resulting in formation of acetyl-CoA, including acetyl-CoA, propionyl-CoA, succinyl-CoA and malonyl-CoA. (A) Subcellular sources of acetyl-CoA in eukaryotic cells. Acetyl-CoA is produced within the different subcellular compartments by the action of enzymes, such as acetyl-CoA-synthetases AceS1/2, by fatty acid degradation via β-oxidation in the mitochondrial matrix. Acyl-CoA is produced by the action of long-chain acyl-CoA synthetase (LACS) at the outer mitochondrial membrane. Acyl-CoAs can be transported into and from the mitochondria to the cytosol via the enzyme carnitine palmitoyl transferase I (CPTI). CPTII produces acyl-CoA in the mitochondria. Pyruvate, generated in the cytosol through glycolysis, can be transported into the mitochondria by the unspecific porin VDAC (voltage-dependent anion channel; outer mitochondrial membrane) and the mitochondrial pyruvate carrier (MPC; inner mitochondrial membrane). In the matrix pyruvate dehydrogenase (PDH) oxidatively decarboxylates pyruvate to form acetyl-CoA that can enter the tricarboxylic acid (TCA) cycle. Acetyl-CoA can be transported from the mitochondria into the cytosol in form of citrate by the citrate-carrier (CIC; is part of citrate malate shuttle that exchanges cytosolic malate for mitochondrial citrate). Citrate can be converted to acetyl-CoA by the ATP-citrate lyase (ACL) in the nucleus/cytosol. (B) Biosynthesis of malonyl-CoA. Malonyl-CoA is generated in the cytosol in the process of synthesis of fatty acids from citrate via formation of acetyl-CoA by ACL and conversion of acetyl-CoA by acetyl-CoA-carboxylase to form malonyl-CoA. (C) Biosynthesis of propionyl-CoA and succinyl-CoA. Propionyl-CoA and succinyl-CoA are generated during degradation of branched amino acids, odd chain fatty acids and cholesterol in the mitochondrial matrix. Propionyl-CoA carboxylase produces D-methylmalonyl-CoA from propionyl-CoA. This is converted to L-methylmalonyl-CoA by methylmalonyl-CoA epimerase (MCEE). Finally, succinyl-CoA is formed by methylmalonyl-CoA mutase. Succinyl-CoA is also formed during TCA cycle by oxidative decarboxylation through α-ketoglutarate dehydrogenase from α-ketoglutarate. (D) Formation of acetyl-CoA and acetyl-phosphate in bacteria. Acetyl-phosphate is the major source for non-enzymatic acetylation of lysines in bacteria. Acetyl-phosphate is generated from acetyl-CoA by the action of phosphotransacetylase (Pta). Acetate-kinase (AckA) catalyzes the formation of ATP and acetate from acetyl-phosphate and ADP. The formed acetate can be converted to acetyl-CoA by acetyl-CoA synthetase (Acs).
Structural studies performed by X-ray crystallography revealed that the hydrophobic substrate binding pocket is capable of accommodating longer acetyl-groups connected to the ε-amino-group of lysine side chains up to myristoyl-groups (Wang et al. 2017a,b). Some SIRT2 inhibitors were capable of inhibiting either its deacetylase (AGK2, SirReal2, TM) or the de-fatty acylase activities (peptide-based inhibitor S2DMi-7) (Supplementary Material Figure 1(F)) (Spiegelman et al. 2018). However, the inhibitor S2DMi-7 showed a low level of selectivity as it inhibits SIRT1 and SIRT3 to a similar extent (Kawaguchi et al. 2019). A recent study applied the PROTAC strategy to completely eliminate SIRT2 from cells thereby abolishing the deacetylase and de-fatty acylase activity (Hong et al. 2020). The fact that deacetylase and long chain de-fatty acylase activities can be discriminated by using different inhibitors suggests that either the enzyme uses two active sites or it uses other catalytic strategies that need to be addressed in future studies (Kudo et al. 2018). Our laboratory showed for the substrate Ran, that SIRT2 is capable of deacetylating the acetyl-group on K38 only if the neighboring K37 was acetylated and suggested a di-acetyl-lysine motif deacetylation mechanism (Knyphausen et al. 2016a,b).

The mitochondrial sirtuins: SIRT3, SIRT4 and SIRT5

SIRT3/4/5 are located in the mitochondrial matrix. They are involved in the regulation of metabolism and in the defense of ROS (Singh et al. 2018). All contain a mitochondrial targeting sequence, which is removed upon translocation into the mitochondrial matrix resulting in their enzymatic activation (van de Ven et al. 2017). While SIRT3 is a potent deacetylase that has also remove longer acyl-chains from lysines, SIRT4 shows a weak general deacetylase activity but is an ADP-ribosyltransferase, lipoamidase and deacetylase (Feldman et al. 2015; Min et al. 2018). SIRT5 shows weak deacetylase activity but it has preferences for negatively charged acylations, such as malonylation and succinylation, occurring at lysine side chains (Figure 1(C)) (Du et al. 2011).

The mitochondrial Sirtuin 3 is a potent lysine deacetylase

Studies with mice revealed that SIRT3 is directly involved in ageing as knock-out mice show age-related disorders, such as cancer, neurodegenerative disorders, metabolic syndrome and cardiac hypertrophy (Ansari et al. 2017; Hirschey et al. 2011a,b; McDonnell et al. 2015). Furthermore, knock-out mice develop pulmonary arterial hypertension and acute kidney injury (Ansari et al. 2017). As these mice cannot regulate the levels of ROS, they suffer from mild endothelial dysfunction. One reported substrate of SIRT3 for decomposition of ROS is the mitochondrial MnSOD/SOD2, which is inactivated upon acetylation (Ansari et al. 2017; Qiu et al. 2010). Our laboratory was not able to show deacetylation of natively-folded MnSOD/SOD2 by SIRT3 in vitro suggesting that it either does not constitute a direct substrate or that other unknown factors are missing that activate SIRT3 to deacetylate MnSOD/SOD2 (Knyphausen et al. 2016a,b). This activity to reduce and thereby detoxify ROS makes SIRT3 a tumor suppressor in various tumor types by deacetylating diverse proteins, by counteracting the Warburg effect, i.e. by decreasing the high glycolysis rate observed in tumor cells even in presence of high oxygen levels, and in defense of ROS (Haigis et al. 2012). Genetic deletion of sirt3 resulted in a systemic increase in acetylation of mitochondrial and non-mitochondrial proteins, while knock-out of SIRT4 and SIRT5 did not grossly affect overall acetylation (Hirschey et al. 2011a,b; Sol et al. 2012). This suggests that SIRT3 acts as major mitochondrial deacetylase. Quantitative mass-spectrometry studies revealed that about 20% of all acetylated mitochondrial proteins are substrates for SIRT3, many of which are involved in metabolic pathways, such as β-oxidation, tricarboxylic acid (TCA) cycle, oxidative phosphorylation and amino acid metabolism (Hebert et al. 2013). Data on SIRT3 expression and activity under metabolic stress, i.e. under conditions of fasting or nutrient abundance, are conflicting and strongly depend on the tissue type analyzed. The main role of SIRT3 is to regulate mitochondrial homeostasis (Marcus and Andrabi 2018). Interestingly, SIRT3 levels correlate with the age of the mice used in the experiments. It was shown that aged mice show an upregulation of SIRT3 expression after 24 h of fasting while younger mice showed a decrease (Edgett et al. 2016; Gudiksen and Pilegaard 2017; Hirschey et al. 2010; Jing et al. 2011, 2013; Li et al. 2016a,b; Qiu et al. 2010). Studies in D. melanogaster, S. cerevisiae and C. elegans showed that an overexpression of SIRT3 resulted in an increase in longevity and in humans, reports suggest that polymorphisms resulting in an increase in SIRT3 expression might be connected to a prolonged lifespan (Albani et al. 2014; Lombard et al. 2007). Conflicting reports exist on the role of SIRT3 in the cytosol/nucleus. SIRT3 is reported to be expressed in a long and a short isoform dependent on the tissue. Both isoforms seem to be active in the mitochondria, while some reports suggest a role in the nucleus (Lombard et al. 2007; Scher et al. 2007; Sundaresan et al. 2008). Other reports question the physiological
significance of the short isoform (Cooper et al. 2009; Jin et al. 2009). It is evident that SIRT3 is mainly targeted to the mitochondria and that the main role is to regulate mitochondrial homeostasis and stress response (Bao et al. 2010). Alterations of acetylation states of nuclear proteins, such as FOXO3 by SIRT3, might be indirect by retrograde signaling from mitochondria to the nucleus as observed for PGC-1α (Yu et al. 2017). PGC-1α is regulated indirectly by SIRT3 through the effect of SIRT3 on AMP-activated protein kinase (AMPK) activity, which activates PGC-1α by phosphorylation and its expression level via CREB phosphorylation (Meng et al. 2019; Palacios et al. 2009; Shi et al. 2005; Sundareshan et al. 2009).

The mitochondrial SIRT4 acts as deacylase and ADP-ribosyltransferase

A similar impact on retrograde signaling was recently reported for SIRT4 regulating the activities of AMPK and TORC1 to maintain a balance between anabolic and catabolic metabolism during different states of nutrient availability (Ho et al. 2013; Shaw et al. 2020). SIRT4 was shown to possess weak deacetylase activity. However, it acts as ADP-ribosyltransferase and as deacylase for lipoyl-, methylglutaryl-, hydroxymethylglutaryl-, and 3-methylglutacanoyl-lysine (Anderson et al. 2017a,b; Mathias et al. 2014; Pannek et al. 2017). The activities and substrates of SIRT4 vary in different tissues and cells. SIRT4 interacts with the mitochondrial dynamin-related GTP-binding protein OPA1, thereby regulating mitochondrial quality control and mitophagy (Lang et al. 2017; MacVicar and Langer 2016). SIRT4 was shown to inhibit insulin secretion in pancreatic β-cells by ADP-ribosylation and thereby inactivating glutamate dehydrogenase (GDH). Moreover, SIRT4 was shown to inhibit fatty acid oxidation (β-oxidation) in muscle and fat cells. In that context, SIRT4 was shown to be a potential deacylase to remove methylglutaryl, hydroxymethylglutaryl and 3-methylglutacanoyl from lysine side chains. It was shown that methylcrotonyl-CoA-carboxylase complex (MCCC), which is involved in catabolism of leucine to acetyl-CoA and acetoacetate, is inactivated by these newly found lysine modifications and SIRT4 diminished their levels. Leucine is an allosteric activator for the GDH. Therefore, SIRT4 inactivates GDH by ADP-ribosylation and increases leucine flux by deacylation and activation of MCCC (Anderson et al. 2017a,b). Moreover, SIRT4 was shown to inhibit fatty acid oxidation in mitochondria by deacylation and inactivation of malonyl-CoA-decarboxylase (MCD) resulting in an increase in malonyl-CoA levels (Laurent et al. 2013).

Malonyl-CoA is a key metabolite that elicits the switch from fat catabolism to fat biosynthesis. Malonyl-CoA is converted to acetyl-CoA by MCD and acetyl-CoA can be converted to malonyl-CoA by the biotin-containing enzyme acetyl-CoA-carboxylase (ACC) (Figure 4(B)) (Bowman and Wolfgang 2019; Laurent et al. 2013). ACC is inhibited by phosphorylation through AMPK, a key regulator of cellular energy homeostasis (Herzig and Shaw 2018). Fatty acids are transported for β-oxidation from the cytosol into the mitochondrial matrix by carnitine palmitoyltransferase 1 (CPT1) (Figure 4(A)). This enzyme is inhibited by elevated malonyl-CoA levels. Under nutrient rich conditions (fed state), malonyl-CoA levels increase and mediate fat synthesis while under nutrient poor conditions (fasted state), malonyl-CoA levels decline and mediate transport of fatty acids into the mitochondria for β-oxidation and energy supply (Figure 4(A)) (Hoppel 1982; Lee et al. 2011; Qu et al. 2016). In the liver, SIRT4 inhibits β-oxidation but mechanistically different from muscle and fat cells (Laurent et al. 2013; Nasrin et al. 2010). The liver is a highly metabolically active organ, which synthesizes fats under nutrient rich conditions (lipogenesis) and activates fatty acid oxidation for decomposition of fats during fasting. Under fasting conditions, the liver supplies ketone bodies, i.e. acetoacetate, β-hydroxybutyrate and acetone, for cellular energy supply of the brain. SIRT4 mediates the response of the liver to fasting by affecting the interaction of SIRT1 and peroxisome proliferation-activated receptor α (PPARα) (Laurent et al. 2013). The interaction of SIRT1 and PPARα results in recruitment of SIRT1 to the PPARα response element (PRE). PGCl-α is another interaction partner and coactivator of PPARα. SIRT1 deacylates and thereby activates PGCl-α, the master regulator of mitochondrial biogenesis activating fatty acid oxidation (Canto and Auwerx 2009; Canto et al. 2009). SIRT4 abolishes this interplay thereby acting as an inhibitor of fatty acid oxidation (Duncan 2011). SIRT4 plays a major role on ATP homeostasis by deacylation of the ATP/ADP translocase 2 (ANT2α), an inner mitochondrial transmembrane protein, which results in the reduction of mitochondrial uncoupling and as a consequence leading to increased ATP production (Min et al. 2018). Thereby, a retrograde signaling pathway from the mitochondria to the nucleus is initiated in which the increase in cellular ATP due to SIRT4 mediated inactivation of ANT2α results in inhibition of the cellular energy sensor AMPK (Min et al. 2018). This example shows, that although SIRT4 has a low deacetylase activity, it still can be of physiological significance on specific substrates such as ANT2α. Under conditions of fasting, i.e. under low ATP levels, AMPK
phosphorylates ACC and inhibits its activity resulting in a decrease in cellular malonyl-CoA levels following to CPT1-mediated translocation of fatty acids into the mitochondria for subsequent energy generation via \( \beta \)-oxidation (Figure 4(A)). Another substrate of AMPK is the transcriptional regulator PGC1-\( \alpha \). AMPK directly phosphorylates and activates PGC-1\( \alpha \) and it furthermore induces PGC-1\( \alpha \) expression indirectly by different mechanisms (Jeon 2016).

Next to the impact of SIRT4 on fatty acid oxidation, it was shown to act as a tumor suppressor protein due to its role on mitochondrial metabolism. Upon DNA-damage glutamine catabolism is inhibited. Glu can be generated by deamination from glutamate by the amidohydrolase glutaminase. The Glu is afterwards converted by glutamate dehydrogenase (GDH) to \( \alpha \)-ketoglutarate, which is a metabolite of the TCA cycle and which is important for tumor cell proliferation. By the inhibitory effect of SIRT4 on GDH, SIRT4 acts as a tumor suppressor protein (Haigis et al. 2006; Min et al. 2018; Zhu et al. 2014).

**The mitochondrial SIRT5 shows deacetylase activity towards negatively charged acyl-lysines**

SIRT5 was originally identified as mitochondrial deacetylase with a broad expression profile, being high in brain, heart, liver, kidney, muscle and testis (Michishita et al. 2005; Nakagawa and Guarente 2009; Nakagawa et al. 2009). Today it is known that knock out of sirt5 is well tolerated in mice under unstressed conditions. SIRT5 plays important regulatory roles by regulating glucose oxidation, formation of ketone bodies, fatty acid oxidation, detoxification of ammonia and defense against ROS (Kumar and Lombard 2018). SIRT5 was shown to play a role in neoplasia acting both as a tumor promoter and suppressor in a context dependent manner (Bringman-Rodenbarger et al. 2018; Kumar and Lombard 2018).

Expression of SIRT5 is regulated by PGC-1\( \alpha \) and by AMPK. Overexpression of PGC-1\( \alpha \) resulted in an increase of SIRT5 mRNA level and AMPK activation by metformin resulted in a decreased SIRT5 expression level (Buler et al. 2014). SIRT5 was initially identified as deacetylase and some substrate proteins were reported to be regulated in their activity by SIRT5-mediated deacetylation. The urea cycle enzyme carbamoyl-synthetase 1 (CPS1), which catalyzes the incorporation of ammonia into bicarbonate, was shown to be activated upon SIRT5 mediated deacetylation (Nakagawa and Guarente 2009; Nakagawa et al. 2009).

Genetic deletion of sirt5 resulted in only minor changes in the overall acetylome, suggesting that its main function might not be a deacetylase (Du et al. 2011; Gertz and Steegborn 2010; Park et al. 2013). Later studies showed that SIRT5 acts as a deacetylase with preferences for negatively charged acylations, such as malonyl-, succinyl- and glutaryl-groups (Bringman-Rodenbarger et al. 2018; Du et al. 2011). In the next paragraph we describe the cellular origin of these negatively charged acyl-CoA thioesters that can act as donor molecules for enzymatic and non-enzymatic acylation of lysine side chains.

**Cellular origin of negatively charged acyl-CoA thioesters**

These negatively charged acyl-groups found on lysine side chains originate from coenzyme-A (CoA) thioesters malonyl-, succinyl- and glutaryl-CoA. Malonyl-CoA is generated by the biotin-containing enzyme acetyl-CoA-carboxylase (ACC) by ATP-dependent carboxylation of acetyl-CoA (Figure 4(B)). Humans encode two isoforms, a cytosolic/peroxisomal form (ACC1) that is expressed in lipogenic tissues, such as liver and adipose tissues, and a mitochondrial form (ACC2) that is present in liver but is mainly expressed in skeletal muscle and heart, where not much \( de novo \) fatty acid synthesis takes place. ACC1 catalyzes the first rate-limiting step of \( de novo \) fatty acid biosynthesis in the cytosol. The mitochondrial enzyme ACC2 is supposed to be involved in the regulation of fatty acid oxidation by regulating the carnitine palmitoyl transferase 1 (CPT1), the transporter responsible for transport of fatty acids into the mitochondrial matrix (Figure 4(A)). CPT1 is inhibited by malonyl-CoA (Abu-Elheiga et al. 2000). The intracellular concentration of malonyl-CoA depends on different activities. Firstly, it depends on the activity of ACC1/2. AMPK inactivates ACC1/2 by phosphorylation resulting in increased fatty acid oxidation (Lee et al. 2018). Secondly, the enzyme malonyl-CoA-decarboxylase (MCD) can degrade malonyl-CoA to yield acetyl-CoA and carbon dioxide. MCD is mainly expressed in liver, heart and to some degree in other tissues such as kidney. Humans express two isoforms, a long, mitochondrial isoform and a short cytosolic/peroxisomal form (Figure 4(B)). Deacetylation of MCD by SIRT4 in the mitochondria inhibits its activity. Its main task is conversion of malonyl-CoA to acetyl-CoA for fatty acid biosynthesis. Thirdly, malonyl-CoA can directly be used as a substrate for fatty acid synthesis to palmitate by the cytosolic fatty acid synthase complex and for the fatty acid elongation system for synthesis of long chain fatty acids. The intracellular concentrations of malonyl-CoA are reported to be in the low micromolar range. However, these might vary dependent on the tissue, the cellular organelle and the metabolic state (Saggerson 2008). Succinyl-CoA is a metabolite of the tricarboxylic acid cycle in the
mitochondrial matrix and is formed by oxidative decarboxylation catalyzed by α-ketoglutarate dehydrogenase. Subsequently, succinyl-CoA-synthetase converts the high-energy thioester succinyl-CoA to succinate and CoA under generation of GTP. GTP can subsequently be converted to ATP by nucleoside diphosphokinase. Eukaryotes encode a cytosolic succinate-CoA-synthetase that catalyzes the formation of a high-energy thioester succinyl-CoA and ADP from succinate and CoA under consumption of ATP. During oxidation of odd number fatty acids and metabolism of branched amino acids, such as valine and isoleucine, as well as threonine and methionine and cholesterol, propionyl-CoA is generated. Propionyl-CoA is converted to succinyl-CoA by a multi-enzymatic reaction that involves the propionyl-CoA-carboxylase (PCC), which carboxylates propionyl-CoA to D-methyl-malonyl-CoA (Figure 4(C)). This is converted to L-methyl-malonyl-CoA by methyl-malonyl-epimerase (MCEE). Finally, L-methylmalonyl-CoA is converted to succinyl-CoA by L-methylmalonyl-CoA mutase (MCM) (Figure 4(C)). The intracellular concentration of succinyl-CoA depends strongly on the metabolic state and is reported to be in the low micromolar range. Glutaryl-CoA is produced during lysine and Trp catabolism. In this pathway, α-ketoadipate is converted to glutaryl-CoA by a complex consisting of 2-oxoglutarate dehydrogenase (OGDH)/dihydrolipoyllysine-residue succinyltransferase (DLST)/dihydrolipoyl dehydrogenase (LDL). Glutaryl-CoA dehydrogenase (GCDH) can convert glutaryl-CoA to crotonyl-CoA. It was reported that GCDH activity is inhibited by glutarylation and SIRT5-mediated deglutarylation can restore its activity (Bhatt et al. 2020). Recent data on SIRT5 revealed that genomic deletion of sirt5 results in a significant increase in protein succinylation, malonylation and/or glutarylation of nuclear, peroxisomal and cytosolic proteins (Du et al. 2011). These data suggest that SIRT5 shows deacylase activity outside the mitochondrial matrix. One explanation might be that cells express an N-terminal truncated SIRT5 isoform lacking the mitochondrial targeting sequence but containing the full catalytic domain (NCBI Resource Coordinators 2018).

The nuclear sirtuins: SIRT6 and SIRT7

Sirtuin 6 is a tumor suppressor localized in the nucleus

The nuclear SIRT6 came into the focus as it was shown to directly mediate longevity by acting as a tumor suppressor (Kanfi et al. 2012; Sebastian et al. 2012; Tian et al. 2019). It was shown that SIRT6 is associated with chromatin and the major cellular processes regulated by SIRT6 are metabolism, genomic stability and DNA repair. Recent data showed that SIRT6 acts as a sensor for DNA damage directly binding to DNA in a sequence independent mechanism recognizing DNA double-strand breaks and activating DNA damage response (Onn et al. 2020). The nuclear SIRT6 was reported to possess a low deacetylase activity compared to SIRT1/2/3 but it shows a strong deacetylase activity towards long chain lysine fatty acylated substrates. Recent data showed that the deacetylase activity can be activated by fatty acids and small molecules mostly carrying a negatively charged carboxylate (Feldman et al. 2013; Klein and Denu 2020; Klein et al. 2020). As described for SIRT4, SIRT6 can directly deacetylate specific substrates, such as histones H3K9ac and K3K56ac, in vivo thereby regulating genome stability and gene expression (Gil et al. 2013). This might be due to the fact that the SIRT6 deacetylase activity is activated in vivo by a yet unknown mechanism. This could include activation by free fatty acids, such as myristic acid, or by binding to other proteins, such as lamins (Ghosh et al. 2015; Klein and Denu 2020; Klein et al. 2020). SIRT6 was shown to promote longevity by regulating multiple pathways. It acts as mono-ADP-ribosyltransferase, which is required for PARP-1 activation and by promoting this activation SIRT6 supports DNA repair. SIRT6 was the first chromatin factor that reverses the Warburg effect in colon cancer thereby acting as a tumor suppressor (Sebastian et al. 2012; Warburg 1925). Today it is believed that cancer cells make this metabolic adjustment to enable fast ATP synthesis and to facilitate uptake and incorporation of nutrients into biomass needed in fast proliferating cancer cells (Liberti and Locasale 2016; Vander Heiden et al. 2009). SIRT6 catalyzes the removal of longer acyl-chains acting as fatty acid deacylase, as reported for TNFα and R-Ras2 (Klein et al. 2020; Zhang et al. 2017). This activity was reported to be activatable by long-chain fatty acids. Notably, SIRT6 was shown to be activated by fatty acids, such as myristic, oleic and linoleic acid, in vitro at physiological concentrations suggesting a feedback regulatory mechanism (Feldman et al. 2013).

As SIRT6 was shown to act as a tumor suppressor and it has beneficial effects on the ageing process, modulators of SIRT6 activity were of huge interest for therapeutic applications. Recently, the laboratory of Steegborn reported the small molecule quercetin to act as a low potency SIRT6 activator by binding to an isoform specific binding site, while it weakly inhibited SIRT1-3 and SIRT5 by binding to an alternative binding site at the active site entrance (Supplementary Material Figure 1(G)) (Huang et al. 2018a,b; You and Steegborn 2018, 2020; You et al. 2017,
In contrast quercetin derivatives such as catechin gallate inhibited all sirtuins including SIRT6 (Supplementary Material Figure 1(G)). Future studies will show if endogenous cellular molecules exist that could act as modulators for SIRT6 activity.

**SIRT7 is enriched in nucleoli and can be activated by DNA/RNA**

SIRT7 is localized within the nucleus where it is enriched in the nucleoli associated with rRNA genes and promoter regions. RNA polymerase I and the transcription factor UBF were among the first interaction partners and substrates identified (Chen et al. 2013). SIRT7 was shown to be involved in regulation of rRNA transcription, cell proliferation and apoptosis. Genomic deletion of sirt7 resulted in reduced lifespan in mice and a reduced stress resistance. Next to histones, several non-histone proteins were reported substrates of SIRT7 deacetylase activity, including p53, PAF53, GABP-β1, NPM1 and U3-55k. Recently, SIRT7 was shown to possess lysine-desuccinylase activity for histones with implications on chromatin condensation and DNA double-strand repair (Li et al. 2016a,b). Reports on SIRT7 regulation reveal that SIRT7 activity is stimulated by dsDNA and rRNA/tRNA in vitro (Tong et al. 2016, 2017). In that context it was shown that dsDNA induces SIRT7 activity to deacetylate H3K18ac, and tRNA/rRNA induces the activity of SIRT7 to remove butyryl-, octanoyl- and myristoyl-acyl-groups from H3K9. Further studies revealed that SIRT7 binds to rRNA under physiological conditions using the N-terminal residue preceding the catalytic domain and a polybasic region at the far C-terminus (Tong et al. 2017).

**Sirtuin catalyzed deac(et)ylation**

**The sirtuin structure**

Sirtuins are structurally composed of a Rossmann-fold domain, which is a known dinucleotide binding domain and a Zn\(^{2+}\)-binding domain that is important for structural integrity but not directly involved in catalysis (Figure 3(C)). The Zn\(^{2+}\)-binding domain is composed of a three-stranded antiparallel β-sheet with a varying number of α-helices. Zn\(^{2+}\)-ion binding is achieved by two pairs of strictly conserved cysteine residues. The Rossmann-fold domain consists of a central β-sheet composed of six parallel β-strands that are connected and surrounded by a varying numbers of α-helices depending on the sirtuin protein (Figure 3(C)). Both domains are connected via the co-factor binding loop, that is directly involved in NAD\(^{+}\)-binding. This loop is highly flexible in the NAD\(^{+}\)-free state but it becomes ordered upon NAD\(^{+}\)-binding. The C-site/pocket is formed by residues of the co-factor binding loop that bind the nicotinamide moiety of NAD\(^{+}\). Upon binding to NAD\(^{+}\) the sirtuin adopts an open conformation. NAD\(^{+}\) binds in an inverted conformation compared with other NAD\(^{+}\)-binding proteins: the adenine base binds to the C-terminal half of the β-sheet and the nicotinamide binds to the N-terminal (Sanders et al. 2010). The NAD\(^{+}\)-binding site can be subdivided into three regions, the A-site binding to adenine-ribose, the B-site binding to nicotinamide-ribose and the C-site, the nicotinamide binding site. The phosphates are bound by a conserved Gly-X-Gly motif. Binding of a substrate is mediated by residues on the surface of the sirtuin.

**The catalytic mechanism exerted by sirtuins**

Sirtuins act as NAD\(^{+}\)-dependent deacetylases/deacylases to remove acyl-chain types from lysine side chains dependent on the sirtuin isoform (Figure 3). Moreover, SIRT6 was reported to possess mono-ADP-ribosylation (MAR) activity and for other sirtuins, such as SIRT2 and SIRT6, MAR activity was reported (Frye 1999; Liszt et al. 2005). However, deacetylase activity is several orders of magnitude faster compared to the MAR activity, suggesting that the MAR activity is either of low physiological significance or other factors are missing that stimulate MAR activity of respective sirtuins (Kowieski et al. 2008). Both, deacetylation and MAR depend on the co-substrate NAD\(^{+}\) used in a stoichiometric manner during catalysis. During catalysis the carbonyl oxygen of the substrate acyl-group acts as a nucleophile to attack the electrophilic C-1′ of the NAD\(^{+}\)-ribose resulting in fast release of nicotinamide and formation of a C-1′-O-alkylamidate intermediate (Figure 3(B)). Upon formation of this intermediate, the sirtuin adopts a closed conformation. The alkylamidate is resolved by attack of the C-2′ hydroxyl of the NAD\(^{+}\) ribose, that is activated by a conserved histidine residue, on the O-alkylamidate carbon, resulting in formation of a cyclic 1′,2′-intermediate. An active site water molecule is activated by a general base. This attacks the cyclic intermediate resulting in formation of 2′-O-acetyl-ADP-ribose and the deac(et)ylated lysine (Feldman et al. 2012). The mechanism is still under debate and experimental results support either an S\(_{N}1\), a concerted S\(_{N}2\) or a dissociative S\(_{N}2\)-like mechanism (Figure 3(B)) (Feldman et al. 2012; Sauve 2010). Nicotinamide acts as a non-competitive inhibitor of sirtuins, which can inhibit deac(et)ylation and restores NAD\(^{+}\) by a base-exchange mechanism (Jackson
et al. 2003; Sauve 2010). If nicotinamide is bound in the C-pocket, together with either NAD⁺, ADP-ribose or acetyl-lysine, the conformation of the co-factor binding loop varies slightly. For catalysis, a productive conformation is formed in which the NAD⁺ is compatible with substrate binding and catalysis.

The acetyl-lysine of the substrate binds to the sirtuin in a cleft between the Zn²⁺-binding small domain and the Rossmann-fold domain and protrudes in a hydrophobic tunnel up to the catalytic center. Binding of the substrate to the sirtuin induces significant structural changes in which the Zn²⁺-binding domain is positioned relative to the large Rossmann-fold domain to form the hydrophobic tunnel. Thereby, a closed conformation is formed for catalysis, which is not formed in apo-structures without substrate and NAD⁺.

Kinetic studies revealed that substrate binding precedes binding to NAD⁺ as binding to NAD⁺ is weak in absence of substrate (Feldman et al. 2015). One exception is SIRT6 for which strong binding to NAD⁺ was observed even in absence of substrate. In many crystal structures, the acetyl-lysine containing substrate peptides are bound predominantly by main chain interactions. This points to the fact that the primary sequence of the sirtuin is not the sole determinant of substrate specificity. Other regions of the substrate apart from the residues in direct vicinity to the acetyl-lysine might be important for substrate specificity. Many in vitro biochemical and structural studies only use the isolated catalytic domains and small acetyl-lysine peptides. Future studies will reveal if other parts in the sirtuins located in the N-terminal or C-terminal regions from the catalytic site contribute to specific substrate interactions. Our laboratory solved a crystal structure of SIRT2 in complex with a trifluoroacetylated Ran. This showed that SIRT2 uses two aromatic stacking interactions besides main chain interactions for substrate recognition (Knyphausen et al. 2016a,b). We concluded that this is important for the determination of substrate specificity to favor deacetylation by SIRT2 over SIRT1 or SIRT3. Our studies using natively-folded and site-specifically lysine-acetylated sirtuin substrate proteins revealed that next to the primary sequence the protein structure is a major factor for determination of substrate specificity (Knyphausen et al. 2016a,b). Structures of sirtuins containing regions flanking the catalytic domains in complexes with acetyl lysylated full-length proteins will reveal how sirtuins recognize their substrates and if different sirtuins use different strategies for substrate recognition.

A strategy for designing selective and potent mechanistic sirtuin inhibitors is the use of acetyl-lysine analogs, such as trifluoroacetyl- and/or thioacetyllysine, combined with natural substrate peptides (Figure 1(B)) (Farooqi et al. 2019; Kuhlmann et al. 2017; Smith and Denu 2007a,b; Smith et al. 2008; Spiegelman et al. 2018). These analogs reduce the sirtuin-catalyzed deacetylation rate by several orders of magnitude as shown for trifluoroacetylated peptides (Smith and Denu 2007a,b). This is due to the reduced nucleophilicity at the acetyl groups’ carbonyl oxygen that attacks the C-1’ of the NAD⁺ ribose (Figure 3(B)). For peptides containing thioacetyllysine it was shown that it initially results in fast nicotinamide release but the following steps lead to the formation of a stalled reaction intermediate (Figure 3(B)) (Smith and Denu 2007a,b; Smith et al. 2008). Combining those substrate-based peptides with additional structural components might result in even more potent mechanistic substrate-based inhibitors (Knyphausen et al. 2016a,b; Kuhlmann et al. 2017).

**N- and C-terminal regions in sirtuins modulate their activity and function**

Sirtuins contain a catalytic core domain with a mean size of approximately 275 amino acids, which is flanked by unique N- and C-terminal regions of variable length, depending on the sirtuin protein (Figure 3(A)) (Costantini et al. 2013; Michan and Sinclair 2007). Many structures solved by X-ray crystallography focused on the catalytic core domain as the N- and C-terminal regions of sirtuins often contain regions of high structural flexibility and are intrinsically disordered (Figure 3(A)). However, these regions were found to be important for the regulation of their catalytic activity by autoregulatory mechanisms, by post-translational modifications, for the interaction with other regulatory proteins, for substrate binding and for the determination of substrate specificity, for binding to regulatory molecules as well as for their subcellular localization.

**SIRT1 is regulated by an autoregulatory mechanism including N- and C-terminal domains**

For SIRT1 it was shown by structural and functional studies that the region preceding the catalytic domain, the N-terminal region (NTR), is important for nuclear localization and for STAC (SIRT1 activating compound) binding to increase SIRT1 activity (Figure 3(A)) (Tanno et al. 2007). Moreover, the NTR contributes to the catalytic activity, while the region C-terminal from the catalytic core, the C-terminal region (CTR), was shown to be important for NAD⁺ binding, both domains stimulating SIRT1 deacetylase activity intramolecularly (Pan et al. 2012). Structurally, this CTR extends the six-stranded β-sheet of the Rossmann-
fold domain stabilizing the active site (Cao et al. 2015; Dai et al. 2015). SIRT1 is regulated post-transcriptionally and post-translationally. SIRT1 translation was inhibited by micro-RNAs, such as mi-34a, resulting in the degradation of SIRT1 mRNA. This resulted in an increase in p53 K382-acetylation and thereby activation, and induction of p53 induced apoptosis. At the post-translational level SIRT1 activity is regulated by phosphorylation by various kinases (Flick and Luscher 2012; Pandithage et al. 2008). Casein-kinase 2 phosphorylates SIRT1in the NTR and CTR resulting in an increase in p53 K382 deacetylation and a decrease in apoptosis. SIRT1 was furthermore shown to be methylated by the methyltransferase Set7/9 in its N-terminal region. The interaction of Set7/9 with SIRT1 was shown to inhibit SIRT1 capacity to deacetylate p53 (Liu et al. 2011). SUMOylation within the SIRT1 C-terminus regulates SIRT1 deacetylase activity in response to genotoxic stress (Yang et al. 2007). Recent data suggest that SIRT1, in addition to SIRT2 and SIRT6, but interestingly not the mitochondrial sirtuins SIRT3 and SIRT5, are inhibited by S-nitrosylation occurring on the cysteine thiolates complexing the Zn$^{2+}$-ion in presence of nitric oxide and nitrosothiols (Kalous et al. 2020). SIRT1 was heavily ubiquitinated at several lysines. K48-linked polyubiquitination of SIRT1 results in its proteasomal degradation. An increase in SIRT1 activity was observed upon K734-SUMOylation. Finally, methylation of SIRT1 at several lysine side chains in direct N-terminal vicinity to the catalytic domain was shown to disrupt the interaction with p53.

Next to post-translational modifications, SIRT1 activity is furthermore regulated by endogenous regulators active regulator of sirtuin-1 (AROS) and deleted in breast cancer 1 (DBC-1) (Kim et al. 2007; Zhao et al. 2008). The interactions have structurally not been characterized so far. AROS is a 141 amino acids small nuclear protein, which was shown to enhance SIRT1-mediated p53 deacetylation in vitro and in vivo (Kim et al. 2007). Experimental data and structural modeling suggested, that AROS binds to the NTR resulting in a more compact conformation that is transmitted to the active site modulating SIRT1 activity (Autiero et al. 2008; Lakshminarasimhan et al. 2013; Milne et al. 2007).

Structural data on SIRT1 revealed that small molecules, such as resveratrol, a polyphenol present in red wine, and other synthetic SIRT1-activating compounds (STACs) bind to the NTR preceding the catalytic domain most likely bringing both domains in close proximity inducing conformational changes within the active site resulting in higher enzymatic activity (Supplementary Material Figure 1) (Cao et al. 2015; Dai et al. 2015, 2018; Gertz et al. 2012; Lakshminarasimhan et al. 2013a,b). Several of these synthetic STACs have improved pharmacokinetic properties compared to resveratrol and several clinical trials are ongoing (Supplementary Material Figure I(G)) (Scisciola et al. 2020). The negative regulator of SIRT1, DBC-1 was originally found to be encoded in a genomic region that is homozygously deleted in breast cancer. It was shown that DBC-1 interacts and inhibits SIRT1 activity in vitro and in vivo (Zhao et al. 2008). DBC-1 interacts with the catalytic domain of SIRT1 (Kim et al. 2008). Models suggest that DBC-1 and the CTR have an at least partially overlapping binding site on the catalytic core domain. To this end, binding of DBC-1 to SIRT1 abolishes the intramolecular binding of the activating CTR to the SIRT1 catalytic core, thereby resulting in the observed inhibitory effect (Zhao et al. 2008).

**SIRT2 activity is modulated by the N- and C-terminal regions flanking the active site**

SIRT2 contains short regions N- and C-terminal of the catalytic core region (Figure 3(A)). These regions are mostly dispensable for deacetylase activity but were found to modulate their activities as shown for SIRT1 and SIRT2 (Dai et al. 2015; Li et al. 2015; Pan et al. 2012). Most structural and functional studies performed in vitro used a C- and N-terminally truncated protein. SIRT2 occurs as several distinct isoforms. Isoform 2, lacking the N-terminal 37 residues, is imported into the nucleus but does not associate to chromatin. Isoform 5 lacks the first 76 residues containing the nuclear export signal (NES), i.e. the protein localizes and accumulates in the nucleus. Recent data suggest, that SIRT2 shows basal cytosolic-nuclear shuttling and plays a role within the nucleus by deacetylating histones and other non-histone proteins, such as p53 and p300, upon bacterial infection (Eldridge et al. 2020). In turn, p300 was shown to acetylate SIRT2 thereby switching off its deacetylase activity constituting a regulatory feedback mechanism (Han et al. 2008). While SIRT2 nuclear export was shown to be CRM1 dependent, the import was only recently found to be mediated by importins suggesting that SIRT2 translocation to and from the nucleus is regulated in a RAN-dependent process (Eldridge et al. 2020; North and Verdin 2007a,b). The C-terminal of SIRT2 is intrinsically unfolded and has a negative impact on SIRT2-importin binding and leads to nuclear accumulation. Phosphorylation in the SIRT2 C-terminal by cyclin-dependent kinases was shown to modulate SIRT2 function with consequences on mitosis and on cell motility (North and Verdin 2007a,b; Pandithage et al. 2008).
The mitochondrial sirtuins SIRT3/4/5 contain an N-terminal mitochondrial targeting sequence

The mitochondrial sirtuins SIRT3/4/5 all contain an N-terminal mitochondrial targeting sequence (MTS) that targets the enzymes for translocation to the mitochondrial matrix and it furthermore inactivates the enzymes in the cytosol (Figure 3(A)). SIRT3/4/5 almost exclusively encompass the mitochondrial targeting sequence plus the catalytic core domain (https://www.uniprot.org/). According to the Uniprot database, SIRT4 contains an MTS encompassing the N-terminal 28 residues, for SIRT5 the MTS is located within the N-terminal 36 residues (https://www.uniprot.org/). For human SIRT3 it was reported to contain an N-terminal 25 residue long MTS and that the 142 N-terminal residues are removed upon mitochondrial translocation resulting in a catalytically active 28 kDa protein (Bao et al. 2010; Onyango et al. 2002; Scher et al. 2007; Schwer et al. 2002; Yang et al. 2010). From a structural perspective this is unlikely as the amino acid residues 120–142 form an α-helix that is important for the structural integrity of the Rossmann-fold domain and which is present also in other sirtuins. Moreover, all crystal structures of SIRT3 contain the residues 120–142. Therefore, the exact boundaries of the mitochondrial targeting sequence in SIRT3 must be revisited. Notably, the sequences of mouse and human SIRT3 differ mostly in the N-termini. The N-terminus of mouse SIRT3 is shorter compared to the human enzyme (Yang et al. 2010). This might enable different regulation of enzyme function and activity in mouse and human. SIRT3 almost exclusively consists of the catalytic core without flanking regions. This suggests that the activity is not modulated as observed for other sirtuins, such as SIRT1 and SIRT2, by allosteric or autoinhibitory mechanisms (Cao et al. 2015; Li et al. 2015).

Several metabolic pathways reside within the mitochondrial matrix containing acetyl-CoA concentrations in the millimolar range, which favors together with the alkaline condition non-enzymatic acetylation at lysine side chains (Baeza et al. 2015; Dieterich et al. 2019; Hosp et al. 2017; Pietrocola et al. 2015). It was shown that almost 20% of all mitochondrial acetylated proteins are substrates for SIRT3 suggesting that the activity evolved towards low substrate specificity (Hebert et al. 2013; Marcus and Andrabi 2018).

The information on confirmed post-translational modifications, that are of physiological significance for regulation of SIRT3/4/5 is low. Recently, SIRT3 was shown to be present in a SUMOylated state in the mitochondria abolishing its catalytic deacetylase activity (Wang et al. 2019). During fasting, the SUMO specific protease SENP1 was shown to translocate to the mitochondrial matrix to deSUMOylate and activate SIRT3 resulting in a decrease of mitochondrial acetylation and modulation of metabolism (Wang et al. 2019). One report suggests S-glutathionation of SIRT3 to impair SIRT3 function (Dikalova et al. 2020). A study performed by mass-spectrometry suggests succinylation of SIRT3, while another identified several phosphorylation sites within SIRT3 MTS (Flick and Luscher 2012; Park et al. 2013). If and how this modulates SIRT3 function and/or activity was not investigated.

SIRT6 and SIRT7 N- and C-terminal regions are important for subcellular localization

For SIRT6 it was shown that the C-terminal domain following the catalytic domain is dispensable for its catalytic activity but important for nuclear localization. The N-terminal domain is important for chromatin association (Figure 3(A)). Furthermore, it was shown to be important for full catalytic activity (Tennen et al. 2010). More recent data suggest a multivalent tight interaction of SIRT6 with nucleosomes with contribution of the SIRT6 C-terminus (Klein et al. 2020; Liu et al. 2020).

SIRT7 localizes to the nucleus and more specifically to the nucleolus. A report shows the existence of a cytosolic pool of SIRT7 but it needs to be addressed if this is physiologically significant (Kiran et al. 2013). Nucleolar localization is shown to be dependent on a poly-basic nuclear localization signal (NLS) present in the N-terminus and a poly-basic signal targeting it to the nucleoli present in the C-terminus (Kiran et al. 2013). Recently it was shown that SIRT7 deacetylase and de-fatty acylase activity can be activated by binding to DNA and RNA in vitro. Further studies suggest that rRNA is a physiological activator for SIRT7. It was shown that for SIRT7 activity and its capacity to bind to rRNA the N- and C-terminal regions flanking the catalytic domain are needed (Tong et al. 2017).

Acyl-CoA mediated feedback control of KDACs/Sirtuins

Sirtuins can catalyze a variety of different reactions ranging from deacetylation, ADP-ribosylation and deacetylation of various acyl-groups with different properties (Feldman et al. 2012). As these enzymes are all NAD⁺ dependent enzymes, using NAD⁺ as a stoichiometric co-substrate for catalysis, various functionalities are coupled to the cellular metabolic state (Feldman et al. 2012; Smith and Denu 2006a,b). Substrates can be diverse acyl-groups attached via an amide bond to the ε-amino group of
lysine side chains. The presence of the specific type of lysine acyl-group is a direct measure for the actual cellular metabolic state as these indicate the presence of the respective acyl-CoA donor molecules used for enzymatic or non-enzymatic acylation. The concentrations of various acyl-CoA molecules depend on the cellular metabolic state, the metabolic pathways activated and the tissue analyzed and might fluctuate from very low concentrations to rather high concentrations, which enable enzymatic and even non-enzymatic acylations. For acetyl-CoA it is reported to be present in micromolar concentrations in the nucleus and cytosol, while it can reach millimolar concentrations in the mitochondrial matrix (Pietrocola et al. 2015). Acyl-CoA molecules are reactive CoA-thioesters, which can non-enzymatically modify lysine side chains (Figure 4). The enzymes involved in the production of the respective acyl-CoA molecules are often modified themselves by these acylations, either enzymatically catalyzed by KATs or non-enzymatically due to a local increase in the concentration of the acyl-CoA species (Figure 4(A)). It was found that all enzymes of the major metabolic pathways, such as glycolysis, gluconeogenesis, TCA cycle and fatty acid metabolism, that affect the production or consumption of cellular acyl-CoA precursors for lysine-acylation, are themselves targeted by lysine acetylation (Figure 4) (Choudhary et al. 2009, 2014; Park et al. 2013; Zhao et al. 2010). KDACs/sirtuins that are capable of removing those modifications were found to be functionally implicated in those pathways (Choudhary et al. 2014; Wagner et al. 2017; Zhang et al. 2019a, b). As such, the production of these acyl-CoA molecules can be considered as a feedback regulatory circuit, modulating their own production.

Enzymatic acetylation catalyzed by lysine acetyltransferases (KATs)

Mammals encode a limited number of KATs for acetylation of lysine side chains

The acetylation of lysine side chains is enzymatically catalyzed by lysine acetyltransferases (KATs) (Figure 5). Compared to kinases, approximately 500 being encoded in the human genome, the number of KATs is limited. Today, only 21 acetyltransferases are confirmed KATs in human, while often 22 KATs are listed (Supplementary Material Table 1) (Drazic et al. 2016). The KAT GCN5L1 (alternative name: BLOC1S1) was originally assumed to act as mitochondrial KAT due to its similarity to the nuclear KAT Gcn5. However, later studies showed that it does not have intrinsic KAT activity but that it affects mitochondrial acetylation indirectly (Scott et al. 2018; Wang et al. 2017a, b). The exact number of KATs is controversial and more KATs might be encoded in the human genome or need to be experimentally validated (Montgomery et al. 2014).

In contrast to kinases, KAT substrate specificity is created by the regions adjacent to the CoA-binding site and the catalytic site and by formation of multi-protein complexes that regulate enzymatic activity and substrate recognition (Tsubota et al. 2007). Often, KATs themselves are huge multidomain proteins in which other domains, such as bromo- and chromodomains, are important for subcellular localization. Other acetyltransferases exist, that either are metabolic acetyltransferases, i.e. they acetylate metabolic molecules, or they catalyze the acetylation of the α-amino group acting as N-terminal acetyltransferases. Notably, bacterial and plant lysine acetyltransferases, which belong to the Gcn5-related N-terminal acetyltransferase (GNAT) family, were shown to act on the α-amino groups of proteins (Bienvenut et al. 2020; Christensen et al. 2018). It will be interesting to see if mammalian GNAT members are capable of acetylating the α-amino group of proteins and how their substrate specificity is defined.

Mammalian KATs are classified into at least three major families based on sequence, structure, their catalytic mechanism and their homology to yeast orthologues: Gcn5-related N-acetyltransferases (GNAT), the p300/CREB-binding proteins (P300/CBP) and the MOZ, Ybf2, Sas2, and Tip60 (MYST) family (Figure 5(A)) (Drazic et al. 2016). Other reports classify KATs into five families (Friedmann and Marmorstein 2013; McCullough and Marmorstein 2016). Moreover, additional KATs were identified that cannot be categorized in any of these three families (Supplementary Material Table 1). Those KATs are transcription-factor related, i.e. TAF1/TBP, TFIIIC90 or steroid-receptor co-factors, such as CLOCK, which is involved in regulation of circadian rhythm (Drazic et al. 2016; Friedmann and Marmorstein 2013). All KATs share a structurally related acetyl-CoA binding site (Albaugh et al. 2011; Dutnall et al. 1998a, b; Tan 2001). The adjacent regions can vary considerably in sequence and structure between the families and sometimes even comparing members of the same family (Figure 5(A)). Exceptions are the GNAT family members Gcn5, PCAF and HAT1, which share four regions of sequence homology of about 100 amino acids each (McCullough and Marmorstein 2016; Neuwald and Landsman 1997). Therefore, the identification of novel KATs is demanding and recent reports suggest the existence of orphan KATs encoded in the human genome, which might act as KATs in vivo (Montgomery...
et al. 2014). Further studies are needed to show which of these KATs are of physiological importance. The nomenclature for KATs was unified (Allis et al. 2007). However, often alternative names for KATs are used in the literature. We summarized the KATs including their traditional and novel nomenclature (Supplementary Material Table 1).

Catalytic mechanisms exerted by KATs

The KATs were intensively studied structurally and functionally. For the GNATs and MYST family members, a similar catalytic mechanism was proposed and experimentally validated by structural data and enzyme kinetics (Figure 5(C)) (Berndsen and Denu 2008). In this direct attack mechanism, a ternary complex is formed between KAT, acetyl-CoA and substrate. Afterwards, a catalytic Glu acts as base to deprotonate the substrate lysine, increasing its nucleophilicity. The nucleophilic lysine attacks the electrophilic carbonyl carbon of the acetyl-group in acetyl-CoA resulting in formation of a tetrahedral intermediate. This intermediate collapses to form acetyl-lysine and CoA. (Redrawn and modified from Ali et al. 2018).

Figure 5: Lysine acetyltransferases (KATs) catalyze the transfer of the acetyl-moiety from acetyl-CoA for N-(ε)-lysine acetylation. (A) Structures of selected KATs as representatives of the three KAT families in cartoon representation. upper panel: p300/CBP family: p300 KAT domain in complex with a Lys-CoA bi-substrate inhibitor (PDB: 3B1Y; L1 loop important for inhibitor binding shown in red); middle panel: MYST-family: KAT domain of MOZ in complex with acetyl-CoA (PDB: 2RC4); lower panel: GNAT family: KAT domain of human Gcn5 in complex with acetyl-CoA (PDB: 1Z4R). The closeups show the Lys-CoA bisubstrate inhibitor (upper panel) or acetyl-CoA (middle/lower panel) in ball-and-stick representations. (The figure was generated with PyMOL v.2.3.4 Schrödinger, LLC, New York, NY, USA). (B) Catalytic mechanism exerted by GNAT and MYST family KATs. GNAT and MYST KATs use an active site glutamate as general base to deprotonate the e-amino group of the substrate's lysine. This increases the nucleophilicity of the lysine side chain for attack of the carbonyl carbon of acetyl-CoA. A tetrahedral intermediate is formed that collapses into CoA and the acetylated lysine. (Redrawn and modified from Ali et al. 2018). (C) Catalytic mechanism exerted by p300/CBP KATs. Studies suggest that p300/CBP KATs use a Theorell-Chance hit-and-run-mechanism to catalyze acetyl-transfer from acetyl-CoA to the lysine side chain. This mechanism is characterized by the fact that no stable ternary complex between enzyme, substrate and acetyl-CoA is formed. In this mechanism an active site tryptophan orients the substrate's lysine and lowers the pKa of its side chain for nucleophilic attack of the acetyl-CoA. An active site tyrosine acts as acid to protonate the sulfhydryl group resulting in acetyl-lysine and CoA. (Redrawn and modified from Ali et al. 2018). (D) Non-enzymatic N-(ε)-lysine acetylation. Due to the reactivity of the high energy thioester acetyl-CoA and of the lysine side chain, non-enzymatic acetylation occurs at lysines. This is prevalent in the mitochondrial matrix with a pH 7.8 and concentrations of acetyl-CoA in the millimolar range. Moreover, the reactivity of the side chain of lysines is increased if present in basic patches or if Glu/Asp residues are in a distance supporting deprotonation thereby lowering the lysine side chain's pKa value. The reaction proceeds via formation of a tetrahedral intermediate that decomposes to form acetyl-lysine and CoA. (Redrawn and modified from Ali et al. 2018).
identified. Instead, the backbone amide of a Trp acts as a nucleophile to deprotonate the lysine side chain of the substrate increasing its nucleophilicity (Figure 5(C)). Afterwards, the lysine attacks the carbonyl carbon of the acetyl-group of acetyl-CoA. An active site Tyr protonates the deprotonated sulfhydryl group of CoA following acetyl-group transfer (Figure 5(C)). KATs were initially described as enzymes catalyzing the acetylation of lysines in histone tails affecting the packing of DNA and thereby the gene-expression level (Berndsen and Denu 2008; McCullough and Marmorstein 2016).

**Aspects of subcellular localization and physiological roles of KATs**

Today it is known that KATs have far more substrates than just histones and substrates can be found in all cellular compartments. In fact, HAT1 and HAT4 were shown to be predominantly cytosolic, p300/CBP is mostly nuclear but it has several cytosolic substrates and was shown to translocate from the nucleus into the cytosol upon UV irradiation or upon activation of interferon signaling (Cohen et al. 2004; Kramer et al. 2009; Tang et al. 2007). An investigation of the full substrate range of cytosolic CBP under these interventions has not been done so far. For long time no KAT was identified to be located within the mitochondrial matrix, with the initially suggested KAT GCN5L1 not being confirmed to contain a physiologically important KAT activity (Wang et al. 2017a,b). A recent study showed that the nuclear MYST family KAT MOF can translocate to the mitochondrial matrix in aerobically respiring cells regulating oxidative phosphorylation by controlling expression of mitochondrial genes. This suggests that mitochondrial acetylation is not only non-enzymatic (Chatterjee et al. 2016). Further studies are needed to show if MOF has further substrates within mitochondria under different physiological conditions. Interestingly, the protein Tau (MAPT), involved in the regulation of microtubule dynamics by binding and by stabilizing polymerized microtubules, is hyperacetylated and hyperphosphorylated in Alzheimer’s disease. The acetylation state of Tau (MAPT) is regulated by SIRT1 and p300 (Gorsky et al. 2016; Hwang et al. 2016; Min et al. 2015). Notably, Tau itself was shown to possess an intrinsic lysine acetyltransferase activity and it has not been further investigated if Tau has substrates that might contribute to disease progression (Cohen et al. 2013). Another interesting KAT is the regulator of circadian rhythm CLOCK. The KAT activity of CLOCK was shown to be important to mediate its effects on circadian rhythmicity and activation of CLOCK genes by lysine acetylation of histones and its heterodimeric interaction partner BMAL1 (Doi et al. 2006). Further studies are needed to investigate whether Tau and CLOCK have further substrates and how these contribute to regulate the physiological and pathophysiological roles of Tau and CLOCK.

Bacteria also encode lysine acetyltransferases. While for a long time the GNAT-related KAT PatZ was the only KAT identified in *E. coli*, recent data showed that additional KATs are encoded in the *E. coli* genome. At least four additional KATs, RimI, YiaC, YjaB and PhnO, were experimentally validated to act as KAT acetylating the ε-amino group of lysine side chains. Potential substrates of these four KATs were identified using mass-spectrometry. Future studies will reveal if these are direct substrates of the respective KAT enzyme. Interestingly, for some of these KATs activity towards ε-amino groups was also shown, suggesting that these enzymes have a dual activity using the ε-amino group and the α-amino group as acetyl-group acceptor. This was recently also shown for a plant GNAT enzyme (Bienvenut et al. 2020). If other eukaryotic enzymes, at least of the GNAT family, are capable of catalyzing acetylation of the α-amino-group besides from their primary KAT activity must be investigated (Christensen et al. 2018).

**Acyltransferase activities of KATs**

Several additional acylations next to acetylation were reported to occur at lysine side chains (Figure 1(C)). These acylations include aliphatic acylations, such as propionylation, butyrylation and hydroxybutyrylation, negatively charged acylations, such as malonylation, succinylation, and glutarylation, and fatty acid acylations, such as myristoylation and palmitoylation (Figure 1(C)). In fact, it seems that all acylations produced as activated CoA thioesters during diverse metabolic processes, such as fatty acid metabolism, amino acid metabolism or carbohydrate metabolism, are detectable as modifications at lysine side chains. To judge how and to what extent these modifications play a physiologically important role on regulation of protein function depends on their intracellular concentrations and the availability of enzymes catalyzing the transfer of these acylations to the respective lysine side chain. Members of all three KAT classes, i.e. p300/CBP, GNAT and MYST were shown to act as efficient transferases for acetyl-, propionyl- and butyryl-groups to lysine side chains (Kaczmaraska et al. 2017; Ringel and Wolberger 2016; Simithy et al. 2017). Many KATs show an auto-acyltransferase activity modulating their activities and constituting an acetyl-CoA dependent regulatory feedback mechanism (Hansson et al. 2009; Karanam et al. 2006;
Thompson et al. 2004). For members of the MYST family it was reported that the lysine propionyltransferase activity is comparable to the KAT activity (Han et al. 2018). For the acidic acylations, i.e. malonylation, succinylation, glutaryl- 

exybutyrylation, the KATs are less efficient. Overall it seems that the efficiency of the KATs to use acetyl-CoA as donor molecules decreases with the chain length (acetyl > propionyl > butyryl > malonyl > succinyl > \(\beta\)-hydroxybutyryl > glutaryl > crotonyl). However, for the KAT Tip60 it was observed that it catalyzes butyrylation with higher efficiency followed by succinylation and acetylation showing that its acetyl-chain preference depends on the individual enzyme (Simithy et al. 2017). The KAT p300 was shown to possess lysine crotonyltransferase activity towards histones thereby activating gene expression (Sabari et al. 2015). Other KATs tested do not use crotonyl-CoA as a substrate, suggesting that the planarity and rigidity of C=C-double bond is not favorable for binding to the enzymes’ active sites (Simithy et al. 2017). Recently, another p300 catalyzed acylation, i.e. lysine 2-hydroxyisobutyrylation, was shown to be present at glycolytic enzymes decreasing their activity (Huang et al. 2018a,b). Some proteins, such as small GTP binding proteins of the ARF (ADP-ribosylation factor) subfamily of the Ras superfamily, undergo myristoylation on the N-terminal glycine residue regulating its membrane binding capacity in dependence of its GTP/GDP-loading state (Franco et al. 1996). While this myristoylation on the \(\alpha\)-amino group is irreversible, N-(e)-lysine-myristoylation has been identified to occur on a lysine residue directly next to the N-terminal glycine residue in Arf6, which can be reversed by SIRT2. Both acylations were catalyzed by the enzymes N-myristoyltransferase 1/2 (NMT1/2), which have two GNAT domains (Kosciuk and Lin 2020). Future studies will unravel the distribution of this modification in the proteome. To therapeutically tackle KATs by potent and selective inhibitors has been challenging although it is a promising target as some KATs, such as the MYST KATs KAT6A (MOF) and KAT6B (MORF), were involved in tumorigenesis. For KAT6A and KAT6B recently potent and selective inhibitors with low nanomolar affinity (IC\(_{50}\) KAT6A: 8 nM; KAT6B: 28 nM) were developed by screening a small molecule library and subsequent optimization. These reversible competitive inhibitors compete with acetyl-CoA for binding to KAT6A/KAT6B and they induce senescence and arrested the progression of lymphoma in mice (Baell et al. 2018). Other KAT inhibitors are peptide-CoA conjugates as competitive bi-substrate inhibitors containing a CoA that is covalently linked to Lys of a substrate peptide (Lys-CoA bi-substrate inhibitor; Figure 5(A)). Some of these inhibitors are remarkable in the sense of their inhibitory potency and selectivity. As an example, selective and potent HAT1 bi-substrate inhibitors as peptide-CoA conjugates were developed that show sub-micromolar potency (Ngo et al. 2019).

KATs and sirtuins sense the metabolic state by using ac(etyl)-CoA and NAD\(^+\) as co-factors for catalysis

KATs use ac(etyl)-CoAs as donor molecules for the acylation of lysine side chains and sirtuins use NAD\(^+\) as a stoichiometric co-substrate for the deac(etyl)ation. Thereby, lysine-ac(etyl)ation is a cellular regulatory system by which the cellular metabolic state is sensed and directly translated into the modification of diverse protein functions. This allows to adjust cellular processes to the current metabolic state. As described above, different KATs can use different acyl-CoA molecules as donor molecules for lysine-acetylation.

The intracellular concentrations of various acyl-CoA molecules were shown to be in the micromolar to sub-micromolar range (Simithy et al. 2017). These values were reported for total HeLa cells and their concentrations might fluctuate dependent on the cell and tissue type and the sub-cellular compartment within the cells analyzed. Their concentrations might be significantly higher under certain metabolic states (Simithy et al. 2017). As an example, intracellular concentrations of acetyl-CoA were reported to be approximately 10 \(\mu\)M in the cytosol/nucleus, while it can reach millimolar concentrations in the mitochondrial matrix (Pietrocola et al. 2015).

Acetyl-CoA can diffuse through the nuclear pore, but cannot cross the inner mitochondrial membrane. Cytosolic acetyl-CoA generated in glycolysis can be brought into the mitochondria in the form of pyruvate by the mitochondrial pyruvate carrier (MPC), where it is decarboxylated by the pyruvate dehydrogenase complex (PDH) to form acetyl-CoA, CO\(_2\) and NADH (Figure 4(A)). Acetyl-CoA can be transported from the mitochondria into the cytosol in form of citrate generated in the TCA cycle by the citrate transporter (CIC), which is finally converted into acetyl-CoA and oxaloacetate by ATP citrate lyase (ACL). Propionyl-CoA, butyryl-CoA and succinyl-CoA concentrations were shown to be approximately 1 \(\mu\)M, 0.75 and 0.25 \(\mu\)M in HeLa cells (Simithy et al. 2017), the other acyl-CoA molecules being below these concentrations. As stated above, the concentrations might fluctuate dependent on the metabolic state, the tissue and even the subcellular compartment.
thorough and systematic analysis of concentrations of various acyl-CoA molecules under different physiological conditions in different tissue and cell types and even in different subcellular compartments is a challenging task and has not been done so far. The cytosolic/nuclear concentration of acetyl-CoA is in the same low micromolar range as the \( K_M \) values of KATs suggesting that the activity is linearly correlated to the cellular acetyl-CoA concentration and that KAT activity can be regulated by the abundance of acetyl-CoA (Cai et al. 2011; Lau et al. 2000; Lee et al. 2014). As the activities of KATs decrease with the acyl-CoA chain length in vitro, the \( K_M \) values for these acyl-CoAs are expected to be substantially higher compared to acetyl-CoA. To judge if an enzymatically catalyzed acylation occurs in vivo, the lower cellular concentrations of these alternative acyl-CoAs compared to acetyl-CoA should also additionally be taken into account (Wapenaar et al. 2015). The cellular availability of these acyl-CoA molecules might precisely regulate KAT activities. Performing KAT-catalyzed acylations in vitro using equimolar concentrations of acetyl-CoA and the respective acyl-CoAs showed that most KATs strongly preferred acetyl-CoA to modify histones (Simithy et al. 2017). Notably, most KATs bind CoA and acetyl-CoA and some KATs even bind both molecules with similar affinities enabling a negative feedback mechanism by product inhibition. Moreover, this shows that it is the ratio between acetyl-CoA and CoA rather than the acetyl-CoA concentration alone that determines the activity of KATs. Product inhibition might dominate for the alternative acyl-CoA molecules if not approaching similar high concentrations as acetyl-CoA (Denisov and Sligar 2012; Henry et al. 2015; Pietrocola et al. 2014). These data suggest that for most KATs, acetyl-CoA is the primary substrate but that some KATs additionally have the capacity to use other acyl-CoA thioesters to catalyze acyl-transfer to lysine side chains. In terms of evolution it is speculative to decide if these activities developed as a specialization towards different acyl-chain types or if, in contrast, these KATs show this promiscuity towards diverse acyl-chain length as there is no evolution it is speculative to decide if these activities catalyze acyl-transfer to lysine side chains. In terms of having the capacity to use other acyl-CoA thioesters to the primary substrate but that some KATs additionally determine the activity of KATs. Product inhibition might be generated via the salvage pathway or via the Preiss-Handler pathway from nicotinic acid (Supplementary Material Figure 2). Moreover, NAD\(^+\) can be generated via the salvage pathway starting from nicotinamide or from nicotinamide ribose (Supplementary Material Figure 2). The salvage pathway seems to dominate in immortalized cells (Liu et al. 2018). It depends on the activities of the rate limiting nicotinamide phosphoribosyltransferase (NAMPT) converting nicotinamide (NAM) to nicotinamide mononucleotide (NMN) and nicotinamide mononucleotide adenylyltransferases (NMNATs) transferring AMP to the produced NMN to build NAD\(^+\) (Supplementary Material Figure 2) (Bogan and Brenner 2008; Hopp et al. 2021; Stromland et al. 2019). Several reports show that NAD\(^+\) levels decline during ageing and the administration of NAD\(^+\) or of precursors for NAD\(^+\) biosynthesis, such as nicotinamide riboside (NR) or nicotinamide mononucleotide (NMN), improves health during ageing (Liu et al. 2018; Stromland et al. 2019; Xie et al. 2020). The intracellular concentrations of NAD\(^+\) vary in the different compartments as NAD\(^+\) diffuses through the nuclear pore, NAD\(^+\) concentrations are similar in the cytosol and in the nucleus (Blacker et al. 2014). However, NAD\(^+\) is not able to cross the inner mitochondrial membrane resulting in a separated pool of mitochondrial NAD\(^+\) (Anderson et al. 2017a,b). In the mitochondria NAD\(^+\) is present at approximately 400 \( \mu M \) and in the nucleus/cytosol at about 100 \( \mu M \) (Anderson et al. 2017a,b). The concentrations of free NADH are substantially lower (cytosol: approximately 130 nM; mitochondrial matrix: 30 \( \mu M \)) (Cambonne et al. 2016). The binding of NADH towards sirtuins is substantially weaker compared to NAD\(^+\) suggesting that sirtuin activity is not modulated by the NAD\(^+\)/NADH ratio but by the NAD\(^+\) concentration alone (Anderson et al. 2017a,b; Madsen et al. 2016; Schmidt et al. 2004). For SIRT1-3 \( K_M \) values for NAD\(^+\) were reported to reside in the low micromolar range under conditions of saturated substrate suggesting an ordered binding of substrate followed by NAD\(^+\) (Feldman et al. 2015). For SIRT3 a \( K_M \) value for NAD\(^+\) was reported to be 98 \( \mu M \) in one study, while it was reported to be 880 \( \mu M \) in another study (Feldman et al. 2015; Hirschey 2011). A third study showed NAD\(^+\) affinities for SIRT3 and SIRT5 in the high micromolar range, 710 or 980 \( \mu M \), but this is increased in presence of saturating concentrations of substrate peptide (SIRT3: 260 \( \mu M \); SIRT5: 200 \( \mu M \)) (Fischer et al. 2012). The differences might be reflected by different substrate peptides and concentrations of peptides used in the studies. Measurements of NAD\(^+\) \( K_M \) values at sub-saturating acetyl-lysine substrate concentrations will only provide an apparent \( K_M \) Sirtuins are NAD\(^+\)-dependent deacetylases, i.e. they use NAD\(^+\) as a stoichiometric co-substrate to catalyze lysine deacetylation (Feldman et al. 2012). NAD\(^+\) is an important redox equivalent being reduced to NADH during catabolic pathways, which is used in the mitochondria to drive ATP synthesis via oxidative phosphorylation. These enzymes couple the cellular energetic state to the enzymatic activity translating it to modulate protein function. Sirtuins use NAD\(^+\) to deaclylate substrate proteins. Cellular NAD\(^+\) supply is mediated by either de novo synthesis from Trp, via the salvage pathway or via the Preiss-Handler pathway from nicotinic acid (Supplementary Material Figure 2).
value. The real values depend on the concentration of the substrate in vivo. Without substrate bound, the $K_M$ values are substantially lower for all sirtuins analyzed except for SIRT6 for which $K_M$ values were in the low micromolar range, $(K_M: 27 \mu M)$, even without substrate bound (Feldman et al. 2015). Considering the $K_M$ values of sirtuins for NAD$^+$ residing at a similar level as the cellular concentrations, sirtuin activity can be modulated by fluctuations in NAD$^+$ levels dependent on the cellular metabolic state (Cambronne et al. 2016; Pietrocola et al. 2015; Sivanand et al. 2017; Wellen et al. 2009). As described for acetyl-CoA, the NAD$^+$ concentrations can fluctuate strongly dependent on the metabolic condition, tissue/cell type and the subcellular compartment (Koch-Nolte et al. 2011). At the mechanistic level this is predominantly based on the fact that NAD$^+$ is a co-factor for many enzymes, such as sirtuins, PARPs and CD38/CD157 (Canto et al. 2015). Sirtuins were shown to either directly regulate lifespan in model organisms, such as yeast, flies, worms and mice, or as mediators of dietary restriction to improve health during ageing (Bonkowski and Sinclair 2016; Grabowska et al. 2017; Zullo et al. 2018). Sirtuins mediate this effect by, firstly, modulating the activities of specific substrate proteins by deacetylation and, secondly, by removing systemic acetyl-lysines arising during aging as a kind of detoxifying mechanism. Both mechanisms might be exerted by sirtuins and by different sirtuins to a different extent.

### Non-enzymatic acetylation to control protein function

With the progress in quantitative mass-spectrometry over the last decade lysine acetylation sites were identified in a number that rival phosphorylation (Kouzarides 2000). These data suggest that lysine acetylation is present in all kingdoms of life, in all cellular compartments and in proteins covering all essential cellular functions (Choudhary et al. 2009; Weinert et al. 2013a,b). Data on lysine acetylation and later on other acylations occurring on lysine side chains suggested that alterations in the cellular metabolic state often systemically affected the cellular acylation status.

The global impact on the overall acylation status suggests that the underlying mechanism is a non-enzymatic mechanism rather than a specific mechanism exerted by enzymes (Figure 5(D)) (Baeza et al. 2015). Of course, this does not exclude that under these conditions proteins are altered at individual sites in their acetylation status in terms of a specific enzymatically catalyzed reaction (Baeza et al. 2015; Wagner and Hirschey 2014). The chemistry of acetyl-CoA being a high-energy reactive thioester and that the $\epsilon$-amino group of lysine side chains can be quite reactive, supports the assumption that datasets obtained by quantitative mass-spectrometry often contain non-enzymatic sites. These non-enzymatic reactions occur when acetyl-CoA concentrations are high, as observed within the mitochondrial matrix, in which up to millimolar concentrations of acetyl-CoA were measured and when the reactivity of the lysine side chain is elevated resulting in the increase of its nucleophilicity prone to attack the electrophilic acyl-group of the acetyl-CoA thioester. The increase of lysine side chain reactivity is achieved under more basic conditions within the mitochondrial matrix compared to the cytosol/nucleus (mitochondrial matrix: pH 7.8; cytosol/nucleus: pH 7.4) (Baeza et al. 2016; Porcelli et al. 2005). Furthermore, the primary sequence and the position of the lysine side chain within the structural fold can affect the lysine side chains pKa value and therefore its reactivity (Baeza et al. 2015). The presence of lysine side chains in patches of several basic side chains, referred to as basic patches, lowers their pKa values increasing their nucleophilicity supporting efficient chemical, non-enzymatic acetylation (Baeza et al. 2015; Olia et al. 2015). Moreover, the acidic side chains Glu or Asp in spatial vicinity to a lysine side chain could act as base-like drivers for non-enzymatic acetylation by abstracting a proton from the lysine-side chain thereby increasing the nucleophilicity for attack on the acetyl moiety in acetyl-CoA (Baeza et al. 2015; Drazic et al. 2016; Wagner and Hirschey 2014). Future studies are needed to show how local electrostatics in the primary structure and the three-dimensional structure of the protein affect lysine side chain reactivity (Baeza et al. 2015).

The amount of non-enzymatic acetylation therefore depends on a) the intracellular concentration of acetyl-CoA, b) the primary sequence and c) the tertiary structure of the protein. Importantly, non-enzymatic acetylation can occur as part of tightly controlled and regulated cellular processes, i.e. metabolic fuel switching, and a protein sequence or structure could have been evolved to be prone to non-enzymatic acetylation at a specific lysine side chain as an important regulatory system (Yang et al. 2011). In bacteria non-enzymatic acetylation occurs systemically affecting protein acetylation state. While in eukaryotes acetyl-CoA is the predominant acetyl-donor molecule for non-enzymatic acetylation acetyl-phosphate generated by the reversible actions of acetate-kinase (AckA) and phosphotransacetylase (Pta) during carbon overflow metabolism drives non-enzymatic acetylation in bacteria (Figure 4(A) and (D)) (Weinert et al. 2013a,b). The concentration of acetyl-phosphate in bacteria at steady-state
The readers of ac(et)yl-lysine: bromodomains (BRD) and YEATS domains

Bromodomains

Bromodomains (BRD) were named after the Drosophila protein Brahma, encoded by the brahma gene, in which the BRD was identified first by Tamkun et al. in 1992 (Tamkun et al. 1992). The protein Brahma is part of the Brahma complex involved in chromatin remodeling. Today the BRD is known to be a reader for acetyl-lysine (Figure 1(A), Figure 6(A)) (Dhalluin et al. 1999; Zeng and Zhou 2002). The BRD is a domain composed of approximately 110 residues forming a left-handed four-helix bundle (α, α, α, α) (Figure 6(A)) (Dhalluin et al. 1999; Mujtaba et al. 2007). The loops connecting helices α and α (ZA loop) and helices α and α (BC loop) have a variable length in the different BRDs. These loops form a hydrophobic cavity creating the acetyl-lysine binding site (Figure 6(A)). The N- and C-termini of the BRD are placed on the opposite site of the acetyl-lysine binding site showing that the BRD is an adaptor exposing its acetyl-lysine binding pocket to bind acetyl-lysine on chromatin or other proteins. The binding of a protein depends on the presence of an acetylated lysine side chain, which is bound in a hydrophobic pocket between ZA and BC loops. In several BRDs two Tyr residues (Tyr308 in ZA loop and Tyr350 in α helix in BD2 of mouse BRDT) and one Asn (Asn351 in α helix in BD2 of mouse BRDT) are conserved. The Asn residue (Asn351 in BD2 of mouse BRDT) in the C-terminus of the α helix forms hydrogen bonds with the acetyl-moiety and a Tyr residue (Tyr308 in BD2 of mouse BRDT) in the ZA loop is involved in the establishment of a water network in the acetyl-lysine binding pocket (Figure 6(A)). The surface area surrounding the acetyl-lysine binding pocket in BRDs is highly variable and confers specificity for peptides/proteins. Additional interactions are made with the ZA and/or BC loops and one to two residues neighboring the acetyl-lysine. Finally, residues from ZA and BC loops at the back of the BRD form electrostatic and hydrophobic interactions with residues three or even more positions apart from the acetyl-lysine. Sequence variations in the ZA and BC loops are therefore important for binding specificity. It was shown that a total of 61 BRDs can be found in 46 proteins many of which are chromatin remodeling proteins, such as lysine acetyltransferases, DNA-helicases, and act as transcriptional regulators, while others have functionally not been characterized so far (Filippakopoulos et al. 2012). Notably, recent reports suggest that the BRD reader domains can crosstalk with the catalytic writer domain in KATs within the same polypeptide chain. Along this line the BRD ligand I-CBP112 was found to stimulate p300/CBP KAT activity. This has further consequences for therapeutic applications, such as treatment of prostate cancer (Zucconi et al. 2019). Based on structural alignments, the BRDs can be subdivided into eight families. Several of the 46 nuclear or cytoplasmic BRD containing proteins contain two N-terminal BRDs, namely the bromodomain and extra-terminal domain (BET) proteins BRD2-4 and BRDT. Tandem BRDs are able to act as multivalent cellular recruitment platforms and scaffolds mediated by acetyl-lysine binding resulting in reorganization of chromatin and thereby alteration in gene expression including the enhanced expression of oncogenes, growth-promoting genes and cytokines (Dawson 2017; Gilan et al. 2020; Sharma and Zhou 2015; Tough et al. 2016; White et al. 2019). Moreover, BET proteins were often found to be deregulated in diseases suggesting that they support disease development. Binding of a BRD to its interaction partner can be highly selective. However, binding studies revealed that the interactions are rather weak showing K_D values between 10 and 100 μM (or...
weaker) (Hudson et al. 2000; Mujtaba et al. 2002, 2004, 2007). This enables the design of competitive BRD inhibitors targeting the acetyl-lysine binding site.

BET proteins are promising therapeutic targets in cancer and inflammation. BRD4 was identified as a promising target in cancer and inflammatory diseases as it was found to enhance the expression of hundreds of genes in a cell-type, disease-type and in a context dependent manner (Raux et al. 2016). Pan-BET inhibitors are in clinical trials for the treatment of various cancer types, cardiovascular and inflammatory diseases (Dawson et al. 2011; Mertz et al. 2011; Nicodeme et al. 2010; Roe et al. 2015; Wang et al. 2015; Zuber et al. 2011). BET inhibitors, such as triazolodiazepine JQ1 and its derivatives, are being tested in clinical trials for treatment of various myeloid and lymphoid malignancies, such as NUT midline carcinoma, acute myeloid leukemia (AML), mixed lineage leukemia (MLL) or acute lymphoblastic leukemia (ALL) (Supplementary Material Figure 1) (Dawson et al. 2011; Filippakopoulos et al. 2010; Fish et al. 2012; French 2010; Mertz et al. 2011; Roe et al. 2015; Zuber et al.

Figure 6: The bromodomains (BRD) and YEATS domain are acetyl-lysine reader domains. (A) Structure of mouse BRDT bromodomain 2 (BD2) in complex with acetylated histone H3 peptide. (Upper panel) The bromodomain is an acetyl-lysine reader domain composed of approximately 110 amino acids forming a left-handed helical bundle. (Lower panel) The acetyl-lysine is recognized by interactions of a hydrophobic cavity lined with aromatic residues and by specificity determining interactions, which often involves bridging water molecules. Moreover, a few residues directly in vicinity of the acetyl-lysine mediate binding towards bromodomains as described in the text (PDB code: 2WP1). (The Figure was generated with PyMOL v.2.3.4 Schrödinger, LLC, New York, NY, USA). (B) Structure of GAS41 YEATS domain in complex with a di-acetylated histone H3 peptide (H3K18 and H3K27). The YEATS domain is the second acetyl-lysine reader domain that determines acetyl-group dependent binding to acetylated proteins. (Upper panel) The YEATS domain has an immunoglobulin (IG) fold. The acetyl-lysine is bound in a hydrophobic pocket formed between the β-sheets of the IG domain. (Lower panel) The specificity towards acetyl-lysine is created by several interactions with the acetyl-group, i.e. hydrogen bonds of main chain of Trp92 and side chains of Ser72 and Tyr73. Additionally, several interactions are formed by the YEATS domain with the peptide’s backbone. Interestingly, the GAS41 forms a dimer in which two YEATS domains mediate divalent interactions with two acetyl-lysines within the histone H3 peptide increasing the binding affinity. Moreover, a third YEATS domain molecule makes interactions with the histone H3 peptide not including the acetyl-lysines using the opposite surface of the IG domain. If this is a physiologically important interaction or due to high concentrations in the crystallization drop is not clear (PDB code: 5VNB). (The figure was generated with PyMOL v.2.3.4 Schrödinger, LLC, New York, NY, USA).
However, using these pan-BET inhibitors often results in unwanted off-target effects and toxicity maybe due to the lack of selectivity. To this end, it would be desirable to be able to selectively target the individual BRDs within each BET protein and to specifically target the different BET-proteins (Andrieu et al. 2016). Most of the available BET inhibitors are competitive inhibitors approaching the acetyl-lysine binding site. Progress was made in the efforts to develop more selective inhibitors that target individual BRDs, e.g. BD1 or BD2, in BET-proteins, such as BRD2/BRD3/BRD4 (Cui et al. 2021; Fairev et al. 2020; Gilan et al. 2020; Preston et al. 2020; Watson et al. 2020; Wellaway et al. 2020). To develop potent and selective BET and BRD inhibitors another promising strategy is to target druggable site in BRDs apart from the acetyl-lysine binding site (Olp et al. 2020).

Notably, some BRD are capable of binding to further lysine-acylations. Most human BRDs were shown to bind to shorter lysine-acylations, such as acetyl- and propionyl-lysine (Filippakopoulos et al. 2012; Flynn et al. 2015). The BRDs of BRD9, CECR2 and the second BRD of TAF1 bind crotonyl- and butyryl-lysine in histone derived peptides with similar affinity as acetylated peptides suggesting that these binding capabilities are of physiological importance (Flynn et al. 2015). None of the human BRDs bind to succinyl-lysine containing peptides suggesting that the negative charge interferes with acetyl-lysine binding (Flynn et al. 2015). The exact physiological role of the BRD readers capacity to bind to different acylated lysines and the binding capacity of other BRDs needs further investigation. Some BRDs occur in tandem with a plant homeodomain (PHD)-finger domain allowing multivalent recognition of peptides/proteins with lysine methylation and acetylation (Bienz 2006; Savitsky et al. 2016). These tandem PHD-BRD cassettes were reported to be able to recognize lysine methylation and additional acetylation on histones (Bowket et al. 2018; Savitsky et al. 2016). Thereby, a platform for combinatorial recognition of different PTMs is created that increases the specificity to fully modified proteins as shown for histones (Li et al. 2006; Savitsky et al. 2016; Tsai et al. 2010).

**YEATS (Yaf9, ENL, AF6, Taf14, Sas5)-domains**

Recently, another acetyl-lysine reader domain was identified, the YEATS (Yaf9, ENL, AF6, Taf14, Sas5)-domain (Figures 1(A) and 6(B)). This domain was originally identified in four human proteins and three proteins in S. cerevisiae. All of these proteins are involved in chromatin remodeling, transcriptional regulation or they are components of KAT-complexes. The YEATS domain shows an immunoglobulin (IG) fold consisting of an eight-stranded antiparallel β-sheet. The acetyl-lysine binding is achieved by two loops emanating from the IG domain using a serine-lined aromatic cage where the peptide containing the acetyl-lysine is oriented perpendicular to the β-strands (Li et al. 2014). Comparison of sequences of acetylated histone binding peptides revealed that the acetyl-lysine is mostly within a RKac (Arg-acetyl-lysine) motif most likely accounting for the rather high binding affinities in the range of 4–10 µM observed for YEATS domains for acetylated histone peptides compared to BRDs. The Arg preceding the acetyl-lysine forms a salt bridge with a Glu of the AF9 YEATS domain and the amide NH of the acetyl-lysine side chain is in hydrogen bond distance to a Ser side chain. These interactions result in gain of binding energy and they create specificity in binding (Figure 6(B)) (Li et al. 2014). For the YEATS domain containing proteins AF6 and ENL it was shown that they recruit the histone methyltransferase DOT1L suggesting that the acetyl-lysine binding is used to bring other epigenetic modifiers to chromatin. Recently, mutations within the ENL YEATS domain were identified to occur in Wilms’ tumor and in AML resulting in defects in kidney development and supporting tumorigenesis (Wan et al. 2017). Progress in a small molecule drug discovery assay format resulted in the identification of an ENL YEATS domain acyl-lysine competitive inhibitor with potential applications in treatment of AML (Asiaban et al. 2020). Further studies reported the development of YEATS domain inhibitors for potential therapeutic applications (Asiaban et al. 2020; Garnar-Wortzel et al. 2021; Jiang et al. 2020; Li et al. 2018). Notably, binding of YEATS domains to other lysine-acylations, such as butyryl-, propionyl-, crotonyl-lysine, revealed the general trend that YEATS domains favor binding to crotonyl-lysine over acetyl-lysine by two to sevenfold (Zhao et al. 2016, 2017). Further studies will reveal the physiological significance of binding of YEATS domains to different lysine acylations.

**Role of reader domains in protein recruitment**

The BRD and YEATS domains are used as recruitment modules to proteins that are specifically marked by lysine acylation (Figures 1(A) and 6). This facilitates bringing different enzymes that additionally contain BRD/YEATS domains, such as lysine acetyltransferases (KAT p300/CBP, PCAF TAF1), methyltransferases (MLL, ASH1L), protein kinase (PRKCBP1) and E3-ubiquitin ligase activities (Trim24/28/33/66), to specific intracellular locations that are specified by lysine acylation. Alternatively, BRDs in scaffold proteins (BRD1-4, BRD7, BRD8A/B, BRD9, BRDT,
BRWD3, etc.) can recruit further proteins to specific locations to initiate cellular processes. By BRDs being in tandem with other PTM binding domains, such as PHD domains or chromodomains, to bind methylated lysines, several PTMs can be integrated obtaining combinatorial effects of multivalent PTMs resulting in a fine tuning of the response. Another important aspect of the physiological roles of lysine acylation is inherent to the discovery of acetyl-lysine targeting modules, KATs and KDACs/sirtuins: the establishment of signal transduction cascades based on lysine acylation in a sense of self-amplifying pathways comparable to protein kinase signaling. Individual reports show the existence of these signal cascades based on lysine-acylation. Lysine acetylation was shown to affect interferon signaling. One study suggested that acetylation promotes interferon signaling by recruitment of the KAT CBP to INFαR1 and subsequent acetylation of INFαR2 creating a docking site of IRF9 and the recruitment of STAT1/2 (Tang et al. 2007). However, different studies showed that hyperacetylation by inhibition of KDACs or modulation of KAT activity suppresses interferon signaling (Hansen et al. 2019; Liu et al. 2020; Wieczorek et al. 2012). This suggests that the level of acetylation and the context determines its outcome. Future studies will show to which extent signal transduction pathways are based on lysine acetylation signaling.

Technological advances to study lysine acetylation

Lysine acetylation was identified already in the 1960s to occur on histones regulating RNA synthesis (Allfrey et al. 1964; Phillips 1963). Afterwards, tubulin was one of first cytosolic proteins that was identified to be lysine-acetylated affecting stability of polymerized microtubules (L’Hermault and Rosenbaum 1985; Verdin and Ott 2015). Some advances in tools to study lysine acetylation were important to drive research on lysine acetylation. This includes the development of specific monoclonal antibodies that target acetyl-lysine and which is supposed to discriminate between acetylated and non-acetylated lysine without showing bias in the sequence surrounding the acetyl-lysine (de Boor et al. 2015). Our experience showed that some commercially available acetyl-lysine antibodies were not specific, whereas others are very specific but are not completely unbiased regarding the sequence N- and C-terminal to the acetylated lysine side chain (de Boor et al. 2015; Kremer et al. 2018).

The research on post-translational lysine acetylation was driven by the huge progress in quantitative mass spectrometry allowing to a) identify, b) to assess the dynamics and c) to determine stoichiometries of diverse lysine acylations on a systemic scale. Anti-acetyl-lysine antibodies were used in initial mass-spectrometry based bottom-up approaches to enrich acetylated peptides after tryptic digest. Trypsin cleaves C-terminally of lysine side chains but does not cleave when the lysine side chain is acetylated. To account for the problem of sequence bias of the antibodies mostly commercially available antibodies were mixed for immunoenrichment. Thereby thousands of proteins were identified as lysine-acetylated in human cells (Choudhary et al. 2009; Kim et al. 2006; Lundby et al. 2012). These studies yielded only data on the pure identification of lysine acetylation sites on a site-specific basis but lacked information on its dynamic changes, its regulation or its stoichiometry, which is important to judge the physiological impact of an acetylation site. Afterwards, proteomic workflows were developed based on stable-isotope labeling by amino acids in cell culture (SILAC) enabling to follow relative changes in global acetylation comparing different cellular states (Chen et al. 2013; Choudhary et al. 2014; Ong and Mann 2007; Sol et al. 2012). These studies were performed using diverse conditions, such as genetically deleted or pharmacologically inhibited sirtuins (Chen et al. 2012; Scholz et al. 2015; Sol et al. 2012). Although many of these studies revealed relative changes in overall acetylation, it lacks information of the stoichiometry of acetylation and the results were no proof for identification of direct substrates. Recent studies applied improved workflows including labeling of all accessible non-acetylated lysines by isotopic acetic anhydride or N-acetoxy-succinimide before tryptic digest (Supplementary Material Figure 1(D)). Other workflows combining SILAC and chemical acetylation including validation using internal standards to quantify acetylation stoichiometry (Baenza et al. 2020; Hansen et al. 2019; Lindahl et al. 2019; Weinert et al. 2015). This is a major step to judge the physiological importance of an acetylation site.

For a lysine acetylation creating a loss-of-function on the protein, the modification accumulating to high stoichiometries is a prerequisite while for a gain-of-function low or medium stoichiometries might be sufficient to exert a physiological effect. Overall, these studies revealed that the majority of acetylation sites are of low stoichiometry (<1%) (Weinert et al. 2015). However, several sites showed high stoichiometries of >10% (Baenza et al. 2014; Hansen et al. 2019; Weinert et al. 2015). As stated above, analyzing conditions that favor global acetylation, such as different carbon sources affecting acetyl-CoA levels, will most likely result in alteration of systemic acetylation albeit with rather low stoichiometry while sites that are acetylated by
an enzymatic reaction might yield higher stoichiometries. Using this approach we found that acetylation on histones accumulated to approximately 20% using inhibitors for classical KDACs (unpublished data). Two elegant studies investigated both the kinetics on a short timescale and the stoichiometry of systemic acetylation (Baeza et al. 2020; Weinert et al. 2018). These studies revealed an overlap of dynamic acetylation in different cell lines suggesting similar regulatory control pathways in fundamental processes, such as translation, splicing and proteome homeostasis. Moreover, these studies showed that acetylation stoichiometry is highest in the cellular compartments in which KATs are located, suggesting the major acetylation sites of physiological importance being enzymatically catalyzed rather than non-enzymatically modified sites (Baeza et al. 2020; Weinert et al. 2018). Future studies should focus on conditions that give hints on acetylation sites directly connected to a cellular process and consider that alterations in the acetylation state can be dynamic, transient and fast.

Mass spectrometric analyses to identify substrates often rely on genetic deletion or pharmacological inhibition of KDACs, sirtuins or KATs. While these studies show an alteration of acetylation on a global scale, they do not result in a concise identification of substrates as the observed alterations might be indirect. Therefore, improved experimental setups are needed that allow the identification of direct substrates rather than interaction partners for KATs, KDACs and sirtuins.

Our laboratory uses the genetic code expansion concept (GCEC) to site-specifically incorporate unnatural amino acids into proteins as response to an amber stop codon (Figure 7) (Lammers et al. 2010; Neumann et al. 2008). Using a synthetically evolved aminoacyl-tRNA-synthetase/tRNA_{CUA} pair based on *Methanocaldococcus jannaschii* tyrosyl-tRNA-synthetase/tRNA_{CUA} we incorporate the photocrosslinkers *p*-azido-*l*-phenylalanine (pAz) and *p*-benzoyl-*l*-phenylalanine (pBpa) into the active sites of sirtuins at positions selected by a rational targeted approach (Chin et al. 2002a,b). Thereby, transient enzyme-substrate complexes are stabilized and the covalent substrate-enzyme complexes can afterwards be analyzed by mass spectrometry. Another important methodological development in the research field was the system to genetically encode acetyl-*l*-lysine (Figures 1(B) and 7). This is based on a synthetically evolved acetyl-lysyl-tRNA-synthetase/tRNA_{CUA}-pair from *Methanosarcina barkeri* or *Methanosarcina mazei* originating from the pyrrolysyl-tRNA-synthetase (PylRS)/tRNA_{CUA} (PylT) pair allowing to obtain quantitatively and site-specifically lysine-acetylated proteins (Neumann et al. 2008, 2009). This enables our laboratory and others to study the role of lysine acetylation on protein structure and function and to furthermore use these proteins to study KDAC/sirtuin-catalyzed deacetylation. Using this system we obtained site-specifically lysine-acetylated protein in quantity and quality sufficient to perform biophysical studies including structural analyses using X-ray crystallography. This allowed us to solve the first crystal structures of proteins carrying one or even two acetyl-lysines of structural and functional importance (Kuhlmann et al. 2016a,b; Lammers et al. 2010). These studies revealed that lysine acetylation regulates protein function by various mechanisms apart from neutralizing the positive charge at the lysine side chain. Lysine acetylation affects the interaction with solvent molecules, alters the hydrophobicity, affects the surface-complementarity, it can crosstalk with other PTMs, such as ubiquitination, SUMOylation or methylation, occurring at the same lysine or even on lysines in vicinity, and it affects protein-protein interactions (Chen et al. 2019; de Boor et al. 2015; Knyphausen et al. 2016a,b; Kuhlmann et al. 2016a,b; Lammers et al. 2010). To this end, the effect of a lysine acetylation has to be studied for the individual protein and even site-specifically for the individual lysine acetylation site.

Along this line, using a mimic for lysine acetylation for *in vivo* studies by mutating Lys to Gln to neutralize the positive charge at the lysine side chain or by mutation of Lys to Arg to conserve the non-acetylated, positively charged state can be misleading (Figure 1(B)) (de Boor et al. 2015; Knyphausen et al. 2016a,b). As an example, if the acetylation exerts a steric effect on protein function, replacement of Lys by Arg poorly resembles the non-acetylated state and the Arg would constitute a mimic for the steric components of a lysine acetylation. Similarly, mutation of Lys to Gln, often used to mimic a lysine acetylation, can only be used as reliable mimic for acetyl-lysine if the only mechanism how lysine acetylation affects protein function is charge neutralization rather than exerting a steric effect. Other aspects, such as binding to BRDs or acting as substrates for KDACs and sirtuins, has also not been assessed by these mutants. We observed for some proteins, that Lys to Gln/Arg replacement can be used to study the impact of lysine acetylation on protein function (de Boor et al. 2015). However, for several other cases these mimics are poor tools to study lysine acetylation *in vivo* and the real impact of the acetylation can only be assessed by using the lysine-acetylated protein (Knyphausen et al. 2016a,b). This shows that it has to be decided on the
individual protein and the individual acetylation site, if mimetic mutations are reflecting the real consequences of lysine acetylation.

Notably, using the isolated catalytic domains of SIRT1, SIRT2 and SIRT3 we were not able to reproduce sirtuin catalyzed deacetylation for several reported substrate proteins (Knyphausen et al. 2016a,b). These data on sirtuin-catalyzed deacetylation of site-specifically and natively-folded substrates suggest that the N- and C-terminal domains or a scaffold protein is needed to accomplish substrate deacetylation. Another model is that deacetylation occurs in an unfolded state for example upon translocation into the mitochondrial matrix, where proteins directly encounter conditions that favor non-enzymatic acetylation, i.e. basic pH and millimolar concentrations of acetyl-CoA. These hypotheses need further evaluation.

Variants of *M. barkeri* or *M. mazei* PylRS were developed to incorporate diverse acyl-lysines into proteins, such as lysine propionylation, butyrylation and crotonylation (Figures 1(C) and 7). These systems were used to incorporate acetyl-lysines into histone H3 (Gattner et al. 2013). Further studies to reveal the functional impact of these acylations on protein function and their regulation by KDACs and sirtuins were only marginally performed so far. Of note, systems that allow the incorporation of the acidic lysine-acylations, lysine malonylation, succinylolation or glutarylation are not available today suggesting that the development of these systems face several unresolved problems. It is not clear if the negatively charged acyl-lysines are taken up by the cells, i.e. if they enter the cytosol and/or if these are stable within the cell. Moreover, it is unclear if the step of amino acid activation and charging to tRNA<sub>lys</sub> by the aminoacyl-tRNA synthetases is catalyzed for the negatively charged acyl-lysines. Using the amber suppression system to incorporate acetyl-lysine analogs into proteins other systems based on chemical ligation were developed. However, these systems often are restricted to the far N- and C-termini of proteins (Dawson et al. 1994; Muir et al. 1998).

Several groups reported chemical reactions or the combination of genetically encoding amino acids with chemical reactions applying different strategies to incorporate mono-/di-methyl- and acetyl-lysine into proteins mostly into histones (Chalker et al. 2012; Li et al. 2011; Nguyen et al. 2009, 2010; Simon et al. 2007; Wang et al. 2010). Recently, a chemical approach was reported to specifically incorporate a succinyl-lysine analog into peptides and proteins (Jing et al. 2018). This is a major progress for the study of succinyl-lysine and mechanistic differences to control protein function compared to uncharged and hydrophobic acylations. However, this system has several drawbacks, which makes it difficult for broad application. It relies on modification of a reactive cysteine residue by a thiol-ene reaction with facing problems of specificity if a protein contains several cysteines. Moreover, for use in peptides/proteins a tert-butyl protected ester of N-vinyl-succinamate is used that needs to be deprotected.
Lysines can be modified by acylation, such as malonyl-/glutaryl-lysine, which can be incorporated into proteins. Studies will reveal if other acyl lysines, such as succinyl-/fatty acid lysines, are incorporated into proteins (Murakami et al. 2006; Xiong et al. 2016). Future studies will show if PROTACs can be developed targeting other KATs/KDACs/sirtuins/BRDs using different E3 ligases.

Another important tool to study lysine acylation is the development of artificial flexizymes (Murakami et al. 2006). These de novo ribozymes were used to incorporate acyl tRNAs with a variety of different unnatural amino acids (UNAA). These aminoacylated tRNAs can then be used in vitro translation systems, such as reconstituted E. coli cell-free translation systems, using genetic code reprogramming. This avoids the step of cellular uptake of the acyl-lysine, which might be a critical step particularly using negatively charged acyl-lysines. The basis for the development is that the ribosome seems to accept aminoacyl-tRNAs independent on the amino acid side chain, i.e. these amino acids can be incorporated into the growing polypeptide chain. It was shown that flexizyme accepts a broad variety of different tRNAs and amino acid benzyl esters. The only requirements for the flexizyme to catalyze the aminoacylation of tRNA is the recognition of the 3′-CCA end of the tRNA and that the amino acid contains a benzylic-moiety, such as 3,5-dinitrobenzyl, esterified to the amino acid α-carboxyl group that acts as a leaving group (dFlx: dinitro flexizyme). Using dFlx enables to incorporate acetyl-L-lysine and analogs thereof, such as thioacetyl-lysine, into full length proteins (Murakami et al. 2006; Xiong et al. 2016). Future studies will reveal if other acyl lysines, such as succinyl-/malonyl-/glutaryl-lysine, can be incorporated into proteins using flexizyme. Research in the last decade showed that lysines can be modified by a variety of different acylations, many of which occur on histones and non-histone proteins.

An elegant experimental approach that will help to resolve the contribution of distinct acylations was established by the Neumann group, developing sirtuins with high selectivity towards deacylation of specific lysine-acylations (Spinck et al. 2021). Chemical probes that are based on peptides connected to HDAC-trapping amino acids were developed, which allow the identification of novel interaction partners and substrates (Seidel et al. 2019; Wang and Cole 2020).

A technology that is used today to tackle various cellular targets, such as kinases and nuclear receptors, are the protein targeting chimeras (PROTACs) (Vogelmann et al. 2020). PROTACs are developed to target epigenetic regulators, such as readers, i.e. bromodomains, writers, i.e. KATs, and erasers, i.e. classical KDACs and sirtuins. The underlying principle of PROTACs is the bivalent composition containing a specific ubiquitin E3 ligase moiety and a specific moiety binding the target protein. The E3 ligase results in ubiquitination of the target protein and subsequent proteasomal degradation. The advantage of PROTACs over inhibitors is that PROTACs result in the complete degradation of the target protein, resulting in a longer pharmacodynamic effect alongside with a high level of selectivity and potency. Recent progress was made targeting some KATs, i.e. PCAF and GCN5, BRDs, i.e. BRD7, BRD9 and BET proteins, and KDACs/sirtuins, HDAC6 and SIRT2, using PROTACs (Bassi et al. 2018; Remillard et al. 2017; Schiedel et al. 2018; Testa et al. 2020; Zoppi et al. 2019). However, so far only a limited number of different E3 ligases were applied. Future studies will show if PROTACs can be developed targeting other KATs/KDACs/sirtuins/BRDs using different E3 ligases.

Finally, substantial progress was achieved in protein structure prediction and de novo protein design in recent years (Kuhlman and Bradley 2019). In that context computational approaches were recently applied for de novo design of cyclic peptides targeting protein interfaces that potently and selectively inhibit HDAC6 and HDAC2 (Hosseinzadeh et al. 2021).

Conclusions and perspectives

Lysine acetylation is a very important regulatory system, by which the cellular metabolic state can directly be translated to altered protein activities important for the regulation of essential cellular processes, such as transcription, DNA damage repair, translation and metabolic adjustment to altered conditions. While much research focused on the role of lysine acylation in histones to modulate gene expression, the huge progress in mass spectrometry enabled to show that this PTM is conserved from bacteria to man, it is distributed in all cellular compartments, present in proteins regulating all essential cellular functions. Future studies will reveal further processes and signal transduction pathways that are controlled by lysine acylation. As sirtuins, KATs and KDACs are involved in the development of several severe diseases in a context dependent manner and affect the ageing process, further development of selective and potent activators and inhibitors will be important. Data on the regulation of sirtuins showed that their activities are regulated by diverse mechanisms including autoregulatory mechanisms as observed for SIRT1 and SIRT2, by binding to endogenous regulatory proteins as shown for SIRT1, by...
binding to small molecules as shown for SIRT6, by binding to DNA/RNA as shown for SIRT7 or by post-translational modifications. These studies showed that the deacetylase activity can be stimulated by these mechanisms even of those enzymes that were initially described as weak deacetylases. For many of these regulatory mechanisms the regions flanking the catalytic domain are needed. Future studies will show if activities of other sirtuins are regulated by additional regulatory mechanisms in vivo. The fact that lysine acylation can regulate protein function by various mechanisms necessitates to study its impact on a protein specific and even site-specific basis. The genetic code expansion concept is powerful to study the real consequences of lysine acetylation. In the future it will be important to thoroughly study other lysine acylations and how these are mechanistically different in regulating diverse cellular processes. To this end, the development and improvement of tools to genetically encode other acylations will be valuable. Eligible mass spectrometric workflows allow the determination of absolute quantities of acetylation on a systemic scale. These workflows need to be applied to specific interventions to narrow down which acetylation sites are of physiological significance. The combination of these workflows with other techniques, such as genetically encoding photoactivatable cross-linkers, will result in the identification of novel direct substrates and interaction partners for KATs, KDACs/ sirtuins and for BRDs/YEATS domains. Recently identified enzymes with dual deubiquitinase and acetyltransferase activity injected by Gram-negative pathogenic bacteria into host cells to allow an efficient infection process show that post-translational lysine acetylation is used by bacteria to modulate host cell processes. These enzymes share homology to the known KATs and can acetylate Ser and Thr residues next to Lys. Future studies will show how these bacteria achieve efficient infection and which cellular processes are affected by lysine acetylation. Another important research area that needs to be addressed in the future is the crosstalk with other PTMs, such as ubiquitination, SUMOylation and methylation. As a summary, the research field made remarkable progress in the last years towards understanding post-translational lysine acylation, how it regulates cellular function and how it contributes to disease development. However, many exciting open questions exists that need to be addressed in the future.

Acknowledgments: We thank Dr. Gottfried J. Palm for the helpful discussion of the manuscript.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: This study was funded by the German Research Foundation (DFG; Deutsche Forschungsgemeinschaft) Grant LA2984-5/1 (Project: 389564084) and LA2984-6/1 (Project: 449703098).

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References


Hosp, F., Lassowskat, I., Santoro, V., De Vleeschouwer, D., Fleigner, D., Redestig, H., Mann, M., Christian, S., Hanahan, M., and
The text content is already in a readable and natural format.


Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/hsz-2021-0139).