Mechanisms of histone lysine-modifying enzymes: a computational perspective on the role of the protein environment

Wilian Augusto Cortopassi[a,b], Kiran Kumar[a,b], Fernanda Duarte[a,b], Andre Silva Pimentel[c], Robert S. Paton*[a,b]

[a] Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA, U.K.

[b] Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, U.K.

[c] Departamento de Química, Pontifícia Universidade Católica do Rio de Janeiro, RJ, 22453-900, Brazil.

*robert.paton@chem.ox.ac.uk

Highlights

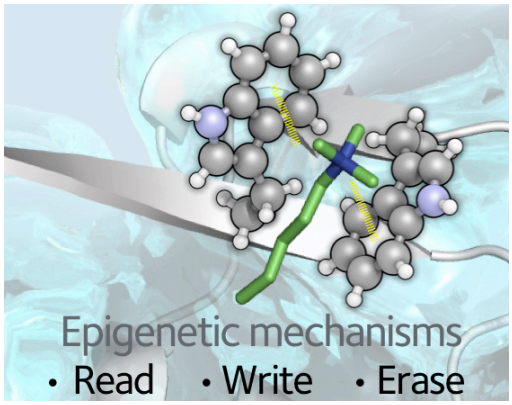
- Overview of epigenetic writing, erasing and reading mechanisms from a chemical perspective.
- Computational modelling is an effective approach for understanding the molecular basis of selectivity and activity of these processes. A literature review summarizing these studies is presented.
- The importance of considering the protein environment when modelling these processes is discussed.
- Hybrid techniques such as QM/MM with the inclusion of dispersion corrections are efficient approaches for determining accurate energetics.
- Long-time MD simulations are necessary to describe conformational changes involved in substrate binding.
- Energy Decomposition Analysis (EDA) is a powerful tool for quantifying the role of specific protein-interactions on histone modifying and reading mechanisms.

Abstract

Epigenetic pathways are involved in a wide range of diseases, including cancer and neurological disorders. Specifically, histone modifying and reading processes are the most broadly studied and are targeted by several licensed drugs. Although there have been significant advances in understanding the mechanistic aspects underlying epigenetic regulation, the development of selective small-molecule inhibitors remains a challenge.

Experimentally, it is generally difficult to elucidate the atomistic basis for substrate recognition, as well as the sequence of events involved in binding and the subsequent chemical processes. In this regard, computational modelling is particularly valuable, since it can provide structural features (including transition state structures along with kinetic and thermodynamic parameters) that enable both qualitative and quantitative evaluation of the mechanistic details involved. Here, we summarize knowledge gained from computational modelling studies elucidating the role of the protein environment in histone-lysine modifying and reading mechanisms. We give a perspective on the importance of calculations to aid and advance the understanding of these processes and for the future development of selective inhibitors for epigenetic regulators.

Graphical Abstract: A tryptophan cage is important for methylated lysine recognition by PHD fingers.



Keywords

Epigenetics; Quantum Mechanics/Molecular Mechanics (QM/MM); Molecular Dynamics (MD); Post Translational Modifications (PTM); Acetyltransferases; Methyltransferases; Bromodomains; Demethylases; Deacetylases; Chromodomains; Tudor Domains; PHD Fingers; Energy Decomposition Analysis (EDA)

1. Introduction

Epigenetic modifications refer to changes in the structure of DNA and the associated histone-proteins, rather than in the DNA sequence itself (i.e. the genetic code). The result of these modifications is the activation or silencing of specific genes. These processes are highly dynamic, involving a number of proteins, which introduce (*epigenetic writers*), recognize (*epigenetic readers*), or remove (*epigenetic erasers)* chemical modifications. Methyl groups are added to the nucleobases of DNA while histones can undergo several modifications including the addition and removal of methyl, acetyl and phosphoryl groups. The downstream effects of these chemical changes impact upon transcription, DNA repair and replication.[1]

The first recognized epigenetic change in humans was DNA methylation, which is involved in transcriptional silencing, gene regulation, development and tumorigenesis.[2] Histone modifications, also known as Post-Translational Modifications (PTM), which broadly describe changes made to proteins following their synthesis by the ribosome, were described later and are among the most studied epigenetic changes.[3, 4] Histones consist of a globular octamer shaped core, around which DNA wraps, and an unstructured tail domain. Most Histone PTMs occur at the N-terminal tails, and less commonly in their globular domains.[5] To date, more than 85,000 Histone PTMs have been experimentally identified; involving the addition of over 25 different chemical groups, such as phosphoryl, acetyl, methyl and hydroxyl.[6] Among them, lysine acetylation/deacetylation and methylation/demethylation have been shown to be the most common.

The emergence of small-molecule inhibitors targeting these epigenetic regulators has stimulated interest in understanding the molecular basis underlying epigenetic regulation, as well as in the design of more potent and selective inhibitors. Over the past few years, an increasing number of these compounds have undergone clinical trials,[1] and some (**Figure 1**) have already been approved for clinical use by the Food And Drug Administration (FDA). For example, cytosine analogues 5-azacytidine and decibitabine block

DNA methylation (i.e. a writing inhibitor) and are used for the treatment of myelodysplastic syndromes,[7, 8] while romidepsin and vorinostat (both targeting epigenetic erasers) are used to treat cutaneous T-cell lymphoma.[9-11] Inhibitors of epigenetic-reader bromodomains are in an early-stage of drug development, including OTX015, currently in the initial phases of clinical trials as a haematological malignancy therapeutic.[12] However, despite these advances, determining at a molecular level how these inhibitors interrupt the native chemical reactions associated with epigenetic modifications remains challenging. A better understanding of the chemical events underlying substrate recognition and epigenetic modifications offers the potential to aid the design of more biologically active compounds with greater target specificity.

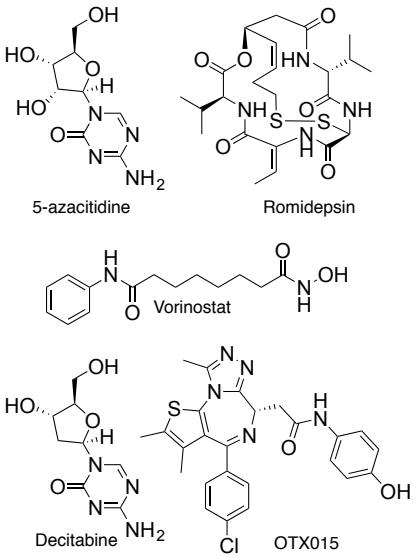
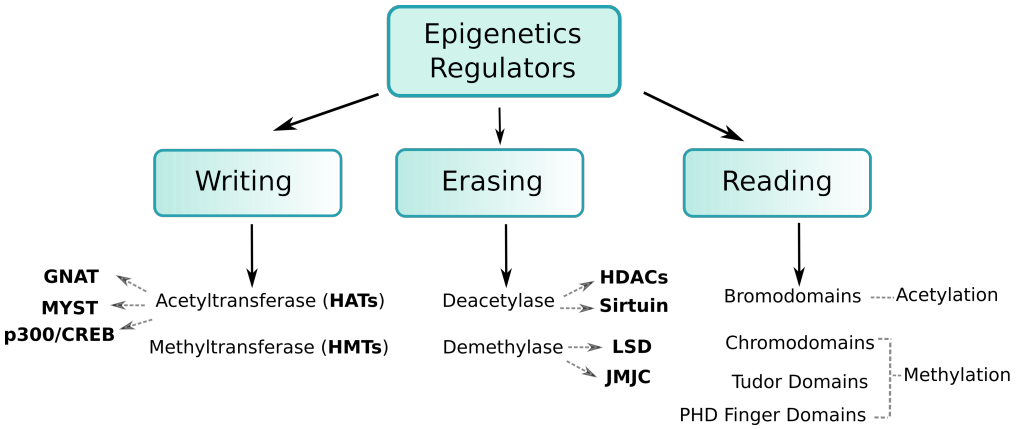


Figure 1. Chemical structures of representative epigenetic FDA-approved drugs: 5-azacytidine, decibitabine, romidepsin, vorinostat and clinical-trial phase compound OTX015.

Computational modelling of epigenetic processes has been a topic of significant research in recent years. Important reviews in the field include Vellore and Baron,[13] who have surveyed the work on histone deacetylases, demethylases and tail dynamics, giving valuable explanations of the background biology in each case; and Smith and Denu (2009),[14] who covered both experimental and computational studies on the chemical

mechanisms of histone lysine and arginine modifications. In the present perspective, we will expand on these works by reviewing recent computational studies, using both Molecular Dynamics (MD) and Quantum Mechanics/Molecular Mechanics (QM/MM) techniques, on histone lysine modifying enzymes. These studies emphasise the role of the protein environment for the binding of reactants and stabilization of transition states and intermediates. Our focus is on mechanistic work rather than non-covalent inhibition: we refer readers to recent excellent reviews concerning the development of epigenetic inhibitors [15-20] and recent works regarding accurate binding free energy predictions.[21-23] A valuable resource of publicly available experimentally determined binding sites of histone tails in complex with human proteins has been developed as an open access web server.[24]

In this perspective, we begin by outlining the key chemical processes regulating histone writing, erasing and reading, as shown in **Scheme 1**. Following this, we will provide a brief overview of relevant computational approaches used to study these processes, highlighting recent advances and challenges. To illustrate this point, we will present key examples in the field. Finally, we conclude by discussing the future role of computational approaches in the continued study of these mechanisms.



Scheme 1. Schematic overview of the classification of epigenetic protein families and their function as presented in the main text.

2. Overview of Epigenetic Regulators

Epigenetic regulators can be divided into three main groups based on their functions: writers, erasers and readers (**Scheme 1**). Additionally, they can be grouped into families – usually structurally related - and according to the chemical process they catalyse. In this section we will briefly introduce the structural, kinetic and chemical features associated with the addition/removal of methyl and acetyl groups to histone lysines. A more detailed description on the chemical aspects of modifications to arginine residues, not discussed here, is presented in ref. [14].

2.1 Epigenetic Writing Regulators

Histone acetylation and methylation are epigenetic writing modifications involved in a large number of druggable targets [19, 25, 26] and have been observed for a wide range of histones, such as the core histones H3, H4 and H2B, dimers containing three alpha helices linked by two loops.[27] The writing process itself refers to the addition of an acetyl or methyl group to the ε-amino group of a lysine at the N-terminal tail of the histone, catalysed by Histone Acetyltransferase (HAT) and Methyltransferases (HKMT) enzymes, respectively. The reaction involves the participation of Acetyl coenzyme A (Ac-CoA) as a cofactor for acetylation, (**Figures 2 and 3**) and S-Adenosyl-L-methionine (AdoMet) as a cofactor for methylation. The reaction mechanism for both processes has been extensively studied experimentally – through structural and kinetics analysis [28-31] – and computationally, as discussed in detail in **Sections 3.1.1** and **3.1.2**.

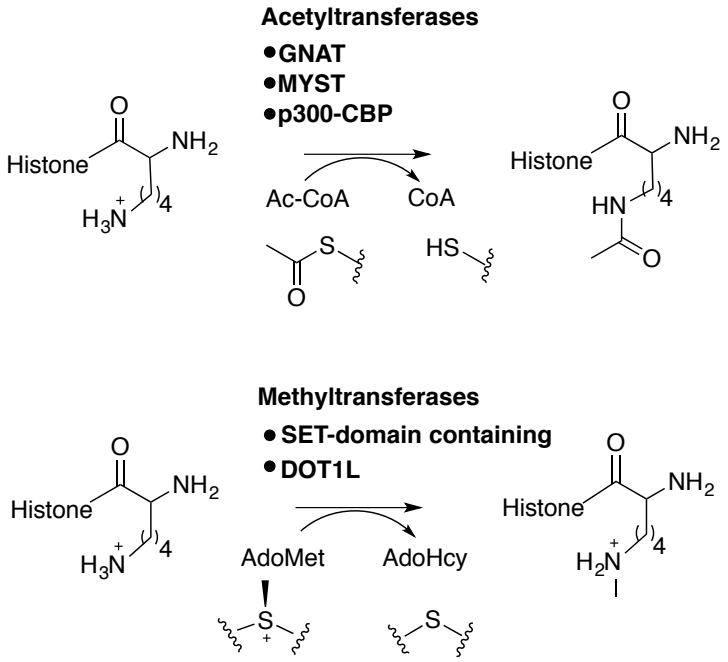


Figure 2. Overview of histone lysine acetylation and methylation mechanisms.

Histone Lysine Acetyltransferases (HATs)*:* Most HATs can be grouped into three families based on sequence, structural, and functional similarities.[32] They include i) the GNAT (GCN5-related N-terminal acetyltransferases) family; (ii) the MYST (Moz, Ybf2/Sas3, Sas2, Tip60) family and (iii) the p300/CREB-binding protein (CBP/CREBBP) family. GNAT proteins contain ~160 residues, forming the HAT domain and a bromodomain, which is responsible the reading/recognition process. Proteins belonging to the MYST family have a ~250 HAT residues, a zinc finger, and an N-terminal chromodomain, which is involved in the reading/recognition process of methylated lysines. Finally, the CBP family proteins are ~500 residues long and contain a HAT-domain, a zinc-finger domain and a bromodomain.[33]

Experimental kinetic data for several HATs suggest the exact sequence of cofactor- and substrate-binding is protein family dependent.[30, 31, 34-36] For example, GNAT proteins are proposed to follow a sequential mechanism **(Figure 3)**,[30] in which following binding of the Ac-CoA cofactor and the substrate in the active site, a glutamate residue activates the positively

charged lysine for its nucleophilic attack at the Ac-CoA carbonyl group. This leads to the formation of a HAT/Ac-CoA/substrate ternary complex. On the other hand, MYST-protein members are proposed to follow a ping pong mechanism, in which the acetyl group is first transferred to a cysteine residue, and then transferred to the substrate.[34] Intriguingly, other structural, kinetics and mutation studies [31, 35] have supported a sequential mechanism for these proteins.

For p300-CBP proteins, it was reported that neither of these mechanisms applies. Instead, a Theorell-Chance, or “hit and run” mechanism has been suggested, where the histone-Ac-CoA-HAT ternary complex never accumulates.[36] Acetyl transfer occurs as an apparent bimolecular collision since the ternary complex is formed so fleetingly that its concentration is always essentially zero, as shown in **Figure 3**.

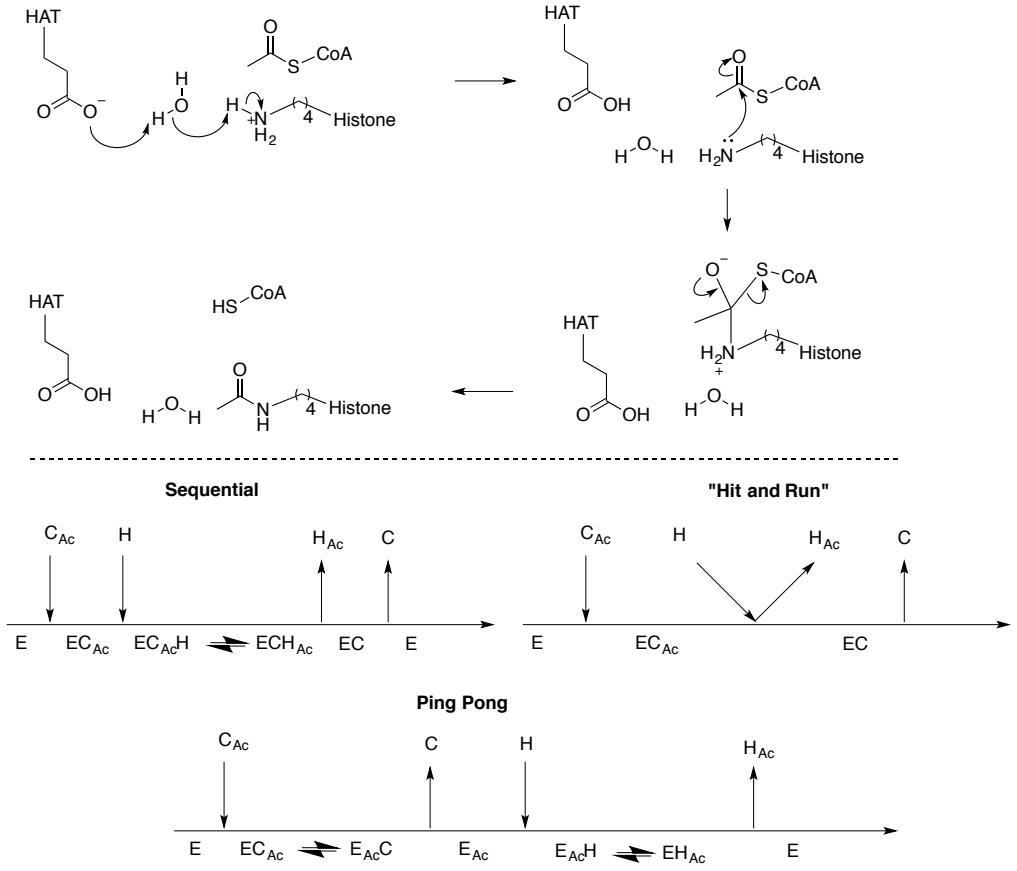


Figure 3. Top: lysine acetylation mechanism in GNAT proteins following a sequential mechanism. Bottom: kinetic mechanisms suggested for acetylases. For the sequential mechanism, the acetylated cofactor (C_{Ac}) and the histone (H) bind to the active site of the enzyme (E) forming a ternary complex ($EC_{Ac}H$). Following the acetyl group (Ac) transfer to the lysine, the acetylated histone (H_{Ac}) and cofactor (C) are then released. In the hit and run mechanism, histone acetylation occurs as a bimolecular step without forming a stable ternary complex. In the ping-pong mechanism, the acetyl group is transferred to the enzyme ($E_{Ac}C$), followed by the release of the cofactor (C). After this, histone binding (forming $E_{AC}H$) and acetylation (forming EH_{Ac}) occur.

Histone Lysine Methyltransferases (HKMTs): HKMTs use the cofactor S-adenosyl-L-methionine (AdoMet) as the methyl donor for lysine methylation. Classification of these proteins relies on the presence or the absence of the Su(var)3-9 Enhancer of Zeste Trithorax (SET) domain, consisting of ~130 residues with a cysteine-rich zinc finger. Most HKMTs contain a SET domain with a few exceptions such as DOT1-Like Histone H3K79 (DOT1L).[37] To date, computational mechanistic studies have exclusively focussed on lysine methylation by SET domain-containing methyltransferases, where the catalytic reaction takes place in these systems. The reaction mechanism involves binding of the AdoMet cofactor, followed by deprotonation of the histone lysine by a tyrosine residue and, finally, the transfer of the AdoMet methyl group to lysine via nucleophilic attack.[37] DOT1L is believed to have a similar mechanism with the main difference being the lysine deprotonation step (**Figure 4**), discussed further in the **Section 3.1.2**.[38] To our knowledge, no computational study has been performed on the DOT1L methylation mechanism, and therefore we focus on the SET domain-containing proteins.

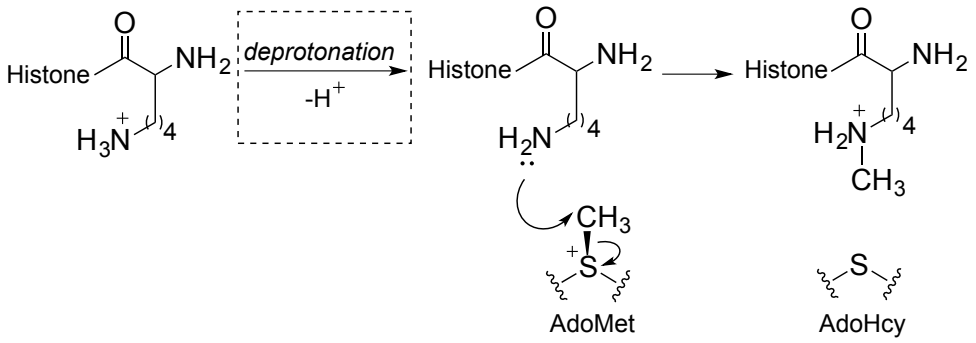


Figure 4. Mechanism of methylation by histone lysine methyltransferases. A number of proposals have been put forward for the first step, which is the activation of lysine, discussed in **Section 3.1.2**.

2.2. Epigenetic Erasing Regulators

Deacetylase and demethylase enzymes perform the removal of acetyl and methyl groups from histone lysines, respectively. The former are already the target of clinically approved small molecules**.**[9-11] These proteins have either metal-dependent active sites or organic cofactors, FAD (Flavin Adenine Dinucleotide) for demethylation and NAD^+ (Nicotinamide Adenine Dinucleotide) for deacetylation. The reaction mechanism for the erasing process is outlined in **Figure 5**, based on experimental and computational studies.[39-44] Due to the relatively recent discovery of these proteins, some of the mechanistic details have been elucidated computationally and await experimental verification.[42, 45] Additionally, as described in Section 3.1.2, computational studies have been essential for understanding active site dynamics, loop flexibility, and binding mechanisms in the erasing regulator proteins.

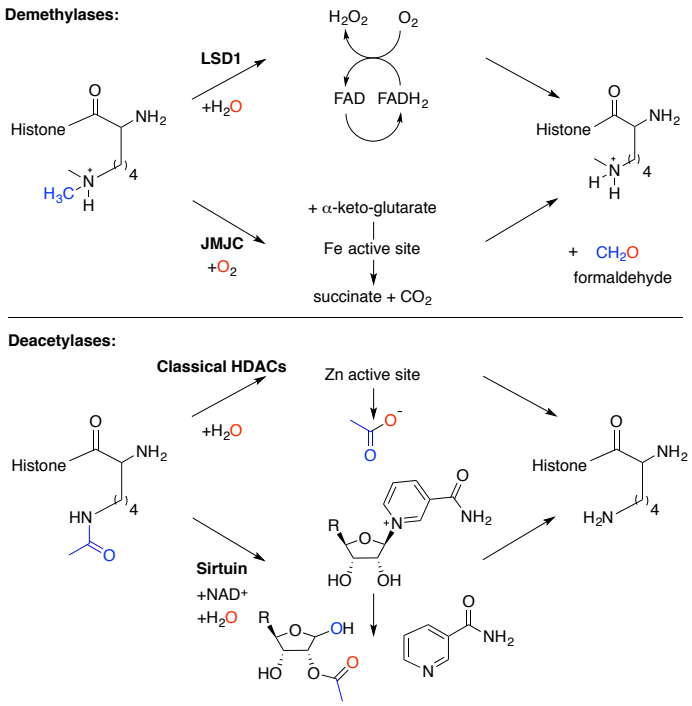


Figure 5. Overview of histone lysine demethylation and deacetylation mechanisms. Demethylases rely on a FAD cofactor, whereas deacetylases require NAD^+.

Histone Lysine Deacetylases **(HDACs):** HDACs are divided into two major families: classical HDACs and Sirtuin. Members of the classical HDAC family are zinc-dependent and share a high-sequence similarity, whereas proteins in the Sirtuin family use NAD^+ as a cofactor. Each of these two major families can be further sub-divided in four classes depending on their structure and function. Class I (HDAC1, HDAC2, HDAC3 and HDAC8), II (HDAC4, HDAC5, HDAC6, HDAC7, HDAC9, HDAC10) and IV (HDAC11) contain proteins that belong to the classical family, while class III (SIRT1, SIRT2, SIRT3, SIRT4, SIRT5, SIRT6, SIRT7, Sir2) belongs to the Sirtuin family.[46, 47] The sequence and cofactor differences in these two families result in different deacetylation mechanisms. Classical HDACs promote the deacetylation of histone acetyl lysines by polarizing the carbonyl group, allowing rate-determining nucleophilic attack at the electrophilic carbon by a water molecule bound to the zinc active site (**Figure 6**). The process is followed by hydrogen abstraction, resulting in the release of the acetyl group of the lysine.[48] In the

Sirtuin family, the mechanism (supported by isotopic and radio-labelling studies) [49] starts with the nucleophilic attack of the acetylated lysine on the ribose ring of NAD^+, forming a positively charged imidate intermediate and releasing nicotinamide. Surrounding active site histidine residues promote the intramolecular nucleophilic attack of the imidate to form a bicyclic intermediate. Hydrolysis results in the net transfer of an acetyl group to the ribose of the cofactor, forming a deacetylated histone.[50]

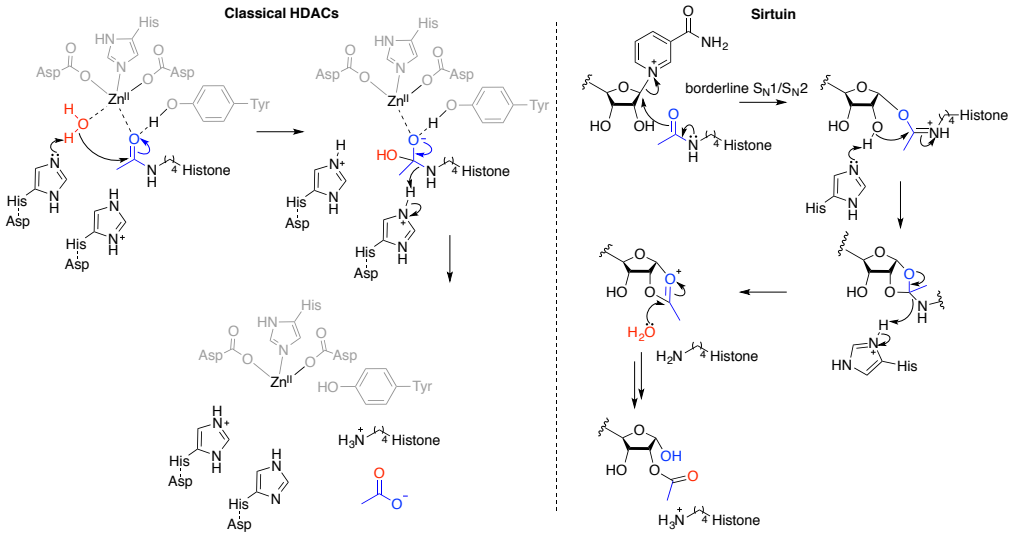


Figure 6. Deacetylation by classical HDACs and Sirtuin family members.

Histone Lysine Demethylases: There are several families of histone demethylases and they can be grouped into two major classes: the amino oxidases Lysine-specific demethylase (LSD), comprising of LSD1 and LSD2, and the Jumonji (JMJC) proteins. LSD proteins are FAD-dependent, whereas Jumonji demethylases have non-heme iron-dependent active sites and a α-ketoglutarate cofactor. Therefore, different mechanisms occur in these proteins (**Figure 7**). In LSD proteins, the reaction begins with the cleavage of the α-CH bond of the methylated lysine in the form of a hydride transfer to the FAD cofactor, converting it to $FADH_2$. The cofactor is then reformed in the presence of O_2, forming H_2O_2. The next step involves the hydrolysis of the imine intermediate with the concomitant release of the demethylated histone and formaldehyde.[51] For JMJC proteins, the first step involves binding of O_2, followed by the oxygen activation, hydrogen abstraction of the methylated

lysine, hydroxyl rebound and the release of the formaldehyde.[52] Due to these structural differences, JMJC proteins are able to catalyze the demethylation of lysine in all possible methylation states (i.e. mono, di- and tri-methylated), whereas LSD proteins are unable to process trimethylated lysine since the iminium intermediate is not viable in this case.

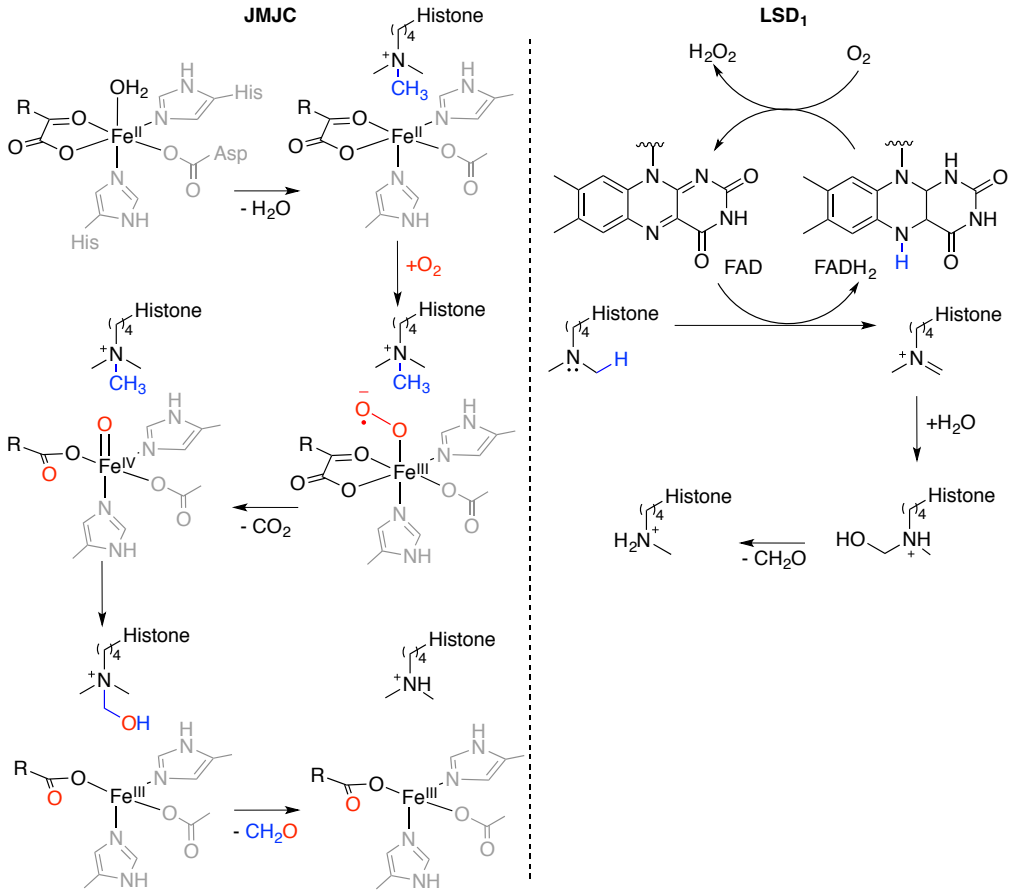


Figure 7. Demethylation by JMJC and LSD proteins.

2.3. Epigenetic Reading Regulators

Following methylation and acetylation, reader proteins can bind and recognize the modified histones to initiate a range of biological pathways, some of which are involved in the progress of diseases, such as cancer and neurological disorders, as well as inflammation and obesity.[53, 54] They can be classified as either acetyl lysine readers, such as bromodomains, or methyl lysine readers, and then subdivided into other classes, such as chromodomains, Tudor domains and The Plant Homeodomain (PHD) fingers (**Figure 8**). Most

computational studies of epigenetic reading proteins cover the role of the protein environment upon substrate binding and recognition. These studies have mainly used classical MD simulations, as there is no bond cleavage/formation in these domains – and therefore a quantum mechanical description of the electronic structure is less important.

A wide array of proteins contains bromodomains, including HATs, transcriptional co-activators (TBP-associated factors (TAFs)), transcriptional mediators (TAF1), and the Bromodomain and Extraterminal domain (BET) family.[17] The bromodomain structure is composed of four helices (αZ, αA, αB, and αC) that are linked by two loop regions (ZA and BC). The binding site includes a hydrophobic pocket and a conserved asparagine residue that directly binds to the acetylated lysine. BET proteins, such as BRD2 and BRD4, can co-operatively bind to two acetylated lysines.[55]

Chromodomains share a similar fold, with three beta sheets packed against a C-terminal alpha helix.[56] However, there are three subdivisions of chromodomains based on their structure: these proteins can have just one chromodomain, paired chromodomains, or similar chromodomains on both N- and C-terminals. Tudor domains usually have around 50 amino acids and are characterized by a bent and anti parallel conformation of beta sheets. The demethylase JMJD2A has two Tudor domains that allow its binding especially to di- and trimethylated lysines at histones H3 and H4. Finally, PHD fingers, protein domains consist of two-stranded beta-sheet and an alpha helix, defining a globular fold.

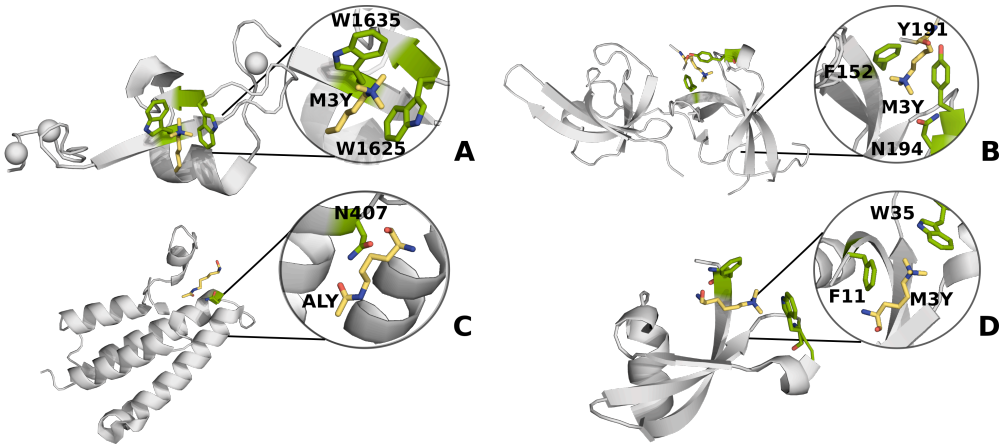


Figure 8. General structures of four major epigenetic reader proteins: A) PHD fingers (PDB ID: 3DB3), B) tudor domains (PDB ID: 3GL6), C) bromodomains (PDB ID: 1E6I) and D) chromodomains (PDB ID: 3I90) shown with select key residues necessary for binding with trimethylated and acetylated lysine histone residues.

3. Computational Methods for Modelling Epigenetic Processes

Over the past decade, molecular modelling has become a well-established tool for understanding chemical and physical properties of proteins involved in epigenetic regulation, and their respective reaction mechanisms. At present, some of these techniques include quantum mechanical (QM) methods and molecular dynamics (MD). In addition, hybrid quantum mechanics/molecular mechanics (QM/MM) approaches have also been used to study reaction pathways within the protein environment.

In general, MD simulations have been used for studying substrate binding in the active site of histone-modifying and reading enzymes. Most of the MD calculations reported here used variations of standard force fields (such as the AMBER ff03 and ff99SB versions) in combination with explicit water (predominantly SPC or TIP3P) models. Most of these studies have been relatively limited in terms of timescale, usually reaching tens of nanoseconds (ns), and relatively few reported studies have been performed at the

microsecond (μs) time scale, which would be needed for exploring larger conformational changes. Notably, Caflisch [57] found that Bromodomain Adjacent To Zinc Finger Domain, 2B (BAZ2B) and CREBBP binding sites were almost completely inaccessible during simulation segments of hundreds of ns, suggesting that simulations on a longer time scale might be necessary to observe an open binding site conformation.

QM/MM studies have mostly used Density Functional Theory (DFT) approaches with the hybrid generalized gradient approximation (GGA) B3LYP functional, in combination with the AMBER force field (we refer to this throughout as B3LYP//AMBER), with relatively few studies including an explicit correction for London dispersion [52] in the QM region. Most of these studies have been done within the ONIOM (Our own N-layered Integrated molecular Orbital and molecular Mechanics) [58] framework.[52, 59-62] Inclusion of an explicit, additive term for dispersion when using B3LYP or other semi-local density functionals (GGAs) provides one way to correct the inadequate description of medium- and long-range electron correlation in the QM region. This impacts upon noncovalent interactions, such as van der Waals forces that will be prevalent for substrate binding and recognition. Grimme's D3 correction provides a density-independent set of atom pairwise dispersion terms, damped at short range, for which global parameters have been obtained for many commonly used exchange-correlation functionals.[63] Small basis sets (e.g. double-zeta) tend to over-stabilize complexes due to the unphysical effect of basis set superposition error, and so dispersion-corrected functionals should be used with larger basis sets to avoid large, unphysical complexation energies. Grimme-D3 dispersion corrections have been shown to have a significant impact on the energy profile of metal-dependent active sites,[62-65] which are relevant to JMJC demethylases and HDACs. The use of modern, dispersion-corrected functionals is also relevant to the study of epigenetically modified DNA sequences, where stacking interactions are prevalent.[66] As an alternative to explicit dispersion correction, exchange-correlation potentials that include a higher derivative of the local density, as in meta-GGA density functionals such as the Minnesota

family (e.g. M06), are also able to capture medium-range correlation resulting in a superior description of noncovalent interactions.

Finally, Energy Decomposition Analysis (EDA) [52, 67] techniques have been a useful tool in both MD and QM/MM studies. These analyses allow one to quantify both van der Waals and electrostatic contributions to the total binding energy due to specific regions of the protein environment. Typically this is done on a per-residue basis. Key residues involved in recognition and catalysis can be identified based on energetic criteria (rather than say, purely by proximity) in the different regulation processes.

3.1 Recent computational studies

In this section we will highlight recent relevant computational studies on the protein families previously described. These studies exemplify both the capabilities of current computational techniques for understanding complex biochemical processes, as well as the current challenges faced in the field.

3.1.1 Mechanistic Insights on Histone Lysine Acetyltransferases

QM/MM studies on lysine acetylation shed light on mechanistic details of the reaction, allowing one to identify the preferred mechanism among a ping pong, sequential or “hit and run” mechanism, as previously discussed in **Section 2.1**. Additionally, MM studies have been performed to evaluate the role of the protein environment on substrate binding/recognition, for example involving dispersion-dominated interactions involving hydrophobic residues and hydrogen bonds involving polar residues. It is important to emphasise that the *trans*-amide configuration of acetyl lysine is energetically preferred (by around 2.5 kcal mol^{-1} according to NMR and computation), such that the occurrence of *cis*-acetyl lysine in protein structures is expected to be rare. However, the existence of *cis*- or twisted acetyl lysine residues in X-ray and NMR structures appears to be erroneously high, and so caution must be exercised in computational studies for structures where this appears to be the case.[68]

GNAT family: As mentioned in **Section 2.1**, experimental data [69, 70] support a sequential mechanism, with the formation of a ternary complex in GNAT proteins. However, the structure of this complex and the catalytic mechanism are not directly accessible from experiment. To investigate these questions, Shen and Luo [61] used MD simulations (using the AMBER ff99 force field in combination with TIP3P explicit solvation) and QM/MM (at the B3LYP//AMBER level of theory) calculations. Using snapshots taken from a 20 ns MD simulation, they modelled the catalytic process by QM/MM calculations. They demonstrated that a glutamate acts as the general base for deprotonating the lysine from H3, a process that is mediated by a water molecule, forming a "proton wire". The subsequent addition of the lysine to the carbonyl of the Ac-CoA was computed to be the rate-determining step. Additional protein-protein interactions were also shown to be important for this process. For example, two GCN5 loops remote from the active site were shown to be crucial for anchoring the substrate – by shaping a cleft where the histone lysine binds – and therefore the subsequent acetylation step.

The acetylation mechanism has also been explored by Jędrzejewski studying a novel human histone acetyltransferase Patt1, that promotes apoptosis in human hepatocellular carcinoma cell lines.[71] Using homology modelling in combination with classical MD simulations (with the AMBER force field and TIP3P explicit solvation), the authors pointed out that a glutamate in Patt1 may act as the general base to activate lysine. The existence of π-π interactions between phenylalanine residues and the aromatic purine ring of Ac-CoA were also found to be relevant for ligand positioning. Additionally, the positioning of a cysteine residue was found to be similar to an analogous cysteine in Esa1 histone acetyltransferase. ESA1 is a member of the MYST family described below, in which the acetyl transfer proceeds through an acetyl-cysteine enzyme intermediate. By analogy, a similar mechanism is plausible for Patt1.

Other groups [72, 73] have studied the complexes formed between the GNT5 proteins GNC5 and p300-CREB binding protein Associated Factor (PCAF)

with the non-histone HIV-1 integrase and the HIV-1 trans-activator of transcription using classical MD techniques. They showed that these complexes can exist in solution, since they were stable along MD simulations using coarse-grained based force fields [72] (200 ns) and Amber [73] (80 ns), and then corroborating experimental data [74, 75]. These studies show that acetylation by GNT5 proteins is also possible in non-histone environments.

MYST Family: Proteins from the MYST family share similar experimental kinetics with GNT5.[14, 76, 77] However, for this family, less is known about the mechanistic details. For Esa1, two mechanisms have been suggested: a sequential one, with the formation of the ternary complex,[31] and a ping pong [34] mechanism, where the acetyl group from the cofactor is first transferred to a cysteine, forming an acetyl-enzyme intermediate, and then to the histone lysine. Few theoretical studies have focused on small molecules binding to MYST proteins. In fact, we have not found theoretical mechanistic studies in support any of either possibility. QM/MM studies comparing both energetic profiles with the glutamate acting as a general base or with the involvement of the cysteine may help in the elucidation of the preferred mechanism.

p300-CBP family: Proteins from the p300-CBP family are suggested to share a unique mechanism known as the “hit and run” mechanism. Unlike GNAT and MYST proteins, no stable ternary complex is formed.[62] The histone associates weakly to the HAT surface, allowing the lysine to go through the p300 tunnel to the active site, with the Ac-CoA already bound. Immediately following acetyl transfer, the acetylated histone is freed.[36] Liu [62] performed 100 ns of MD simulation using the AMBER ff99 force field with SPC explicit solvation and QM/MM with B3LYP//AMBER to investigate the reaction mechanism of p300 acetylation on the H3 histone. Although the substrate is weakly bound to p300, and rapidly dissociates upon acetylation, this step is still necessary for the reaction. This binding is favoured by hydrogen bonds and hydrophobic interactions with the p300 loops. A plausible long-range proton transfer scheme for the first step was shown to have a low energy barrier that was energetically dependent on specific tyrosine,

aspartate and water molecules. Another tyrosine was found to be responsible for the re-protonation of the CoA leaving group to acetylate the histone lysine.

3.1.2 Understanding Histone Lysine Methyltransferase Specificity

One of the controversial aspects of the lysine methylation is how the first step of the reaction, deprotonation of lysine, actually occurs. QM analyses have provided some insights on how the transport of water molecules to and from the active site impacts upon mono- and di-methylation. Other QM studies have focused on the role of key residues responsible for substrate selectivity – it has been computationally and experimentally observed that mutating an active site tyrosine to phenylalanine favour di- over monomethylation.[78, 79]

SET-domain Containing Methyltransferases: Histone Lysine Methyltransferases have been the subject of extensive computational investigations in the last several years.[80-82] Particular attention has been directed towards understanding structural features that are important for the product specificity. Crystal structures[83] have shown that a water molecule occupies the substrate-binding pocket that needs to be released to allow the reaction to proceed. Hu and Yang have studied lysine methyltransferase 9 (SET9) in complex with the tumor suppressor p53 by classical MD simulations using the CHARMM27 force field for ~200 ns.[38] They suggested that experimentally observed dimethylation by SET9 [84, 85] is possible when water molecules, responsible for the lysine deprotonation step, can leave the active site via a channel (**Figure 9**). The channel is formed by a glycine, alanine and two tyrosine residues and remains open when the substrate is bound to SET9 following the first methylation. They compared this with another possible channel, close to Lys372 of p53. However, this second channel is less suitable for dimethylation, since there is inadequate space for water transport when the monomethylated substrate and the cofactor are both bound to the active site.

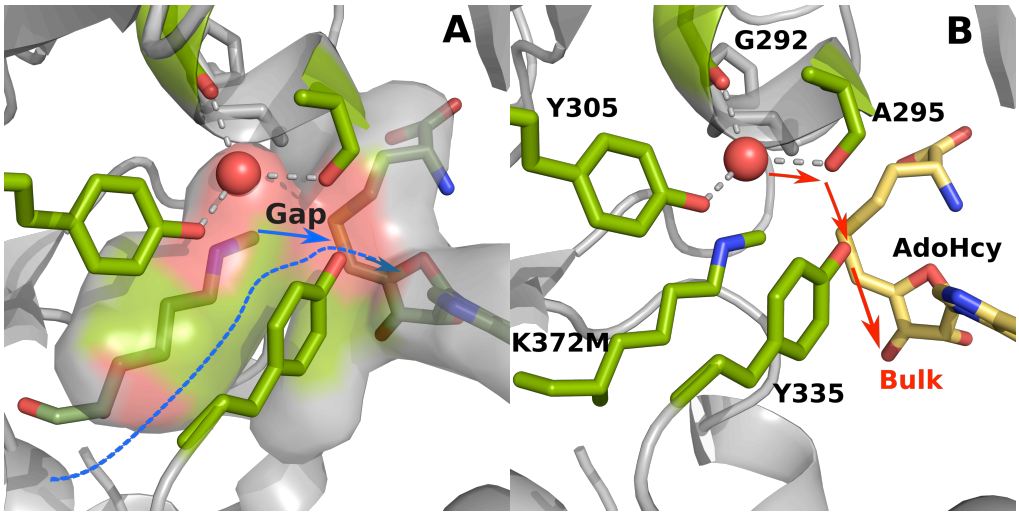


Figure 9. A) and B) are possible water channels in lysine methyltransferase 9. (SET9). Computational work [38] shows that B) is likely to be the preferred channel.

Water channels have also been described as essential for the deprotonation of the histone lysine before undergoing the methylation reaction (**Figure 4)**. While this mechanistic step is still controversial,[14] Zhang and Bruice [86, 87], performed QM/MM calculations with B3LYP//CHARMM31 and found that the AdoMet-SET-histone complex forms an ordered water channel that allows the transport of protons to bulk solvent. For DOT1L, however, the residues located in this region are incapable of performing this deprotonation and there is no evidence suggesting the formation of a water channel similar to the SET-containing proteins. Alternative explanations cite a more hydrophobic active site that reduces the pK_a of the histone lysine and the role of the carboxylate of the AdoMet for the deprotonation of the lysine.[88, 89]

Other studies have focused on how protein conformational changes can regulate the writing process. Xu [90] studied the methyltransferase Nuclear receptor binding SET Domain protein 1 (NSD1) in association with histone H3. They performed 2 ns of MD simulation with the GROMOS96 force field and SPC explicit solvation. In this work, a regulatory loop, which connects the SET and a post-SET domain, was found to be responsible for controlling the degree of lysine methylation, by allowing the access of the histone lysine to

the SET domain of NSD1. This is a unique mechanism in which the size of the entrance site controls specificity for dimethylated lysine.

A combination of QM/MM and MM techniques has greatly aided the investigation of substrate selectivity of SET domain-containing proteins.[91-93] Zhang performed QM/MM Hartree Fock (HF)//Amber calculations finding that the first methyl transfer in SET7/9 enzymes results in a conformational change of the cofactor binding site. This increases the energy barriers for the dimethylation.[94] This methyl transfer in SET7/9 proteins has been proposed to follow a concerted S_N2 mechanism following QM/MM HF//Amber simulations,[91] in which electrostatic contributions from the protein environment lowered the energy barrier for the charge-separated TS structure from 30.9 to 22.5 kcal mol^{-1}.[93] This computed value is in good agreement with the experimental free energy of 20.9 kcal mol^{-1}.[95]

As observed for SET7/9, Guo [96] showed that SET8 also functions preferentially as a monomethyltransferase, rather than a dimethyltransferase, due to the high energetic cost of the substrate reorientation following the first deprotonation step. The possibility of a dimethylation was suggested for the Y334F protein variant, although it was found that a very high activation barrier would deter trimethylation. Guo [97] also analyzed methylation by the Defective In Methylation-5 (DIM-5) protein using QM/MM B3LYP//CHARMM27 calculations. DIM-5 is a trimethylase protein in nature but was observed experimentally to become a mono- or dimethylase following a F281Y mutation. [78] Guo showed that the mutated model destabilizes the methylation transition state, resulting in an 8 kcal mol^{-1} increase in energy barrier for the addition of a third methyl group compared to the first methyl transfer.

Another example of reduced catalytic activity following protein mutation was observed by Chu, Sun and Zhong.[98] They carried out QM/MM simulations on the methyltransferase PR domain zinc finger protein 9 (PRDM9) using B3LYP//CHARMM27 and found that the energy barrier for the first methyl transfer increased following a Y276F mutation, corroborating experiments.[79]

This increase happened because the reactive configuration necessary for the methylation is not observed in the mutated model.

3.1.3 The Role of Protein Dynamics in Histone Lysine Deacetylation

Computational studies exploring epigenetic deacetylation have used MM studies to observe the flexibility of the protein environment upon recognition and binding of reactants. This has been broadly true for both major families, the Classical HDACs and Sirtuin family. Reaction profiles proposed from a QM perspective have also been critical in understanding the role of different classical HDAC histidines as bases and/or proton donors when interacting with the lysine substrates.

Classical HDACs family: Classical histone deacetylases are highly dynamic (notably loop conformations) and a significant number of recent MD studies have been performed to evaluate how conformational changes control the deacetylation process. A particular focus of interest in these systems has been the evaluation of structural and dynamic features involved in the regulation of water access and/or product release. Haider [99] concluded that an arginine residue in HDAC8 acts as a key gate-keeper in this process. They performed 20 ns of MD with the AMBER force field and TIP3P solvation, verifying that this arginine can form hydrogen bonds with the backbone carbonyl group of two conserved glycine residues, blocking one of the HDAC8 channels. Kalyaanamoorthy and Chen have also investigated the gate-keeping process. [100, 101] Using MD simulations and mutagenesis experiments they analyzed the known pathways and also identified additional pathways for inhibitors release and water exchange in HDAC1 and HDAC2 enzymes. Lee [102] performed 5 ns MD with the GROMOS96 force field and the SPC3 solvation to study the dynamics of the tunnel-like active sites of the apoforms of the native and mutated HDAC8, HDAC10, HDAC11. Subtle differences were found in these native and mutated active sites; however, non-conserved residues that are part of the tunnel-like conformation may affect arrangement and dynamics, e.g. a methionine residue in HDAC8.

Arrar [103] showed that the epigenetic regulator inositol tetraphosphate (IP4) plays an important role for the activity of HDAC3. They carried out 100 ns with the AMBER ff99SB force field and TIP3P solvation, finding changes in conformational dynamics were necessary for the activation of HDAC. Interactions with IP4 and the deacetylase-activating domain (DAD) are important in creating an active conformation of HDAC3.

Hunsey and Coveney [104] performed theoretical and experimental studies on the dynamics of the substrate entrance site of HDAC8. Through long-time MD, using the AMBER ff99SB force field with the TIP3P solvation, along with experimental and crystallographic techniques, they found that the substrate entrance is highly dynamic: interactions between two loops (L1, which spans residues 31-35; and L2, residues 83–108) of HDAC8 control enzymatic activity. When these two loops interact with each other, a tyrosine and an aspartate are oriented away from the entrance to the catalytic site. These residues are essential for substrate binding and positioning and the different orientations influence the activity of the enzyme.

Other computational studies on HDACs have shed light on the deacetylation reaction mechanism. Most theoretical QM/MM studies were performed as an attempt to elucidate which of the two active site histidine-aspartate dyads is involved in the process (acting as the base) and which is involved in the proton transfer to the deacetylated histone. Zhang [105] presented an alternative proton-shuttle mechanism for HDAC8, in which a single neutral histidine residue accepts a proton from the active site water molecule and then transfers it to the amide nitrogen, facilitating the release of the acetyl group. However, this alternative mechanism (**Figure 10**) is controversial. Further QM/MM studies by Wiest [59] with an extended active site including aspartates, similar to previous QM studies by Vanommeslaeghe [106] supported the previous model,[107] with one histidine/aspartate dyad functioning as a general base for the reaction and the other, positively charged histidine protonating the epsilon nitrogen atom enabling collapse of the tetrahedral intermediate. These mechanisms were also shown by Gleeson [60] after performing M06//AMBER QM/MM calculations.

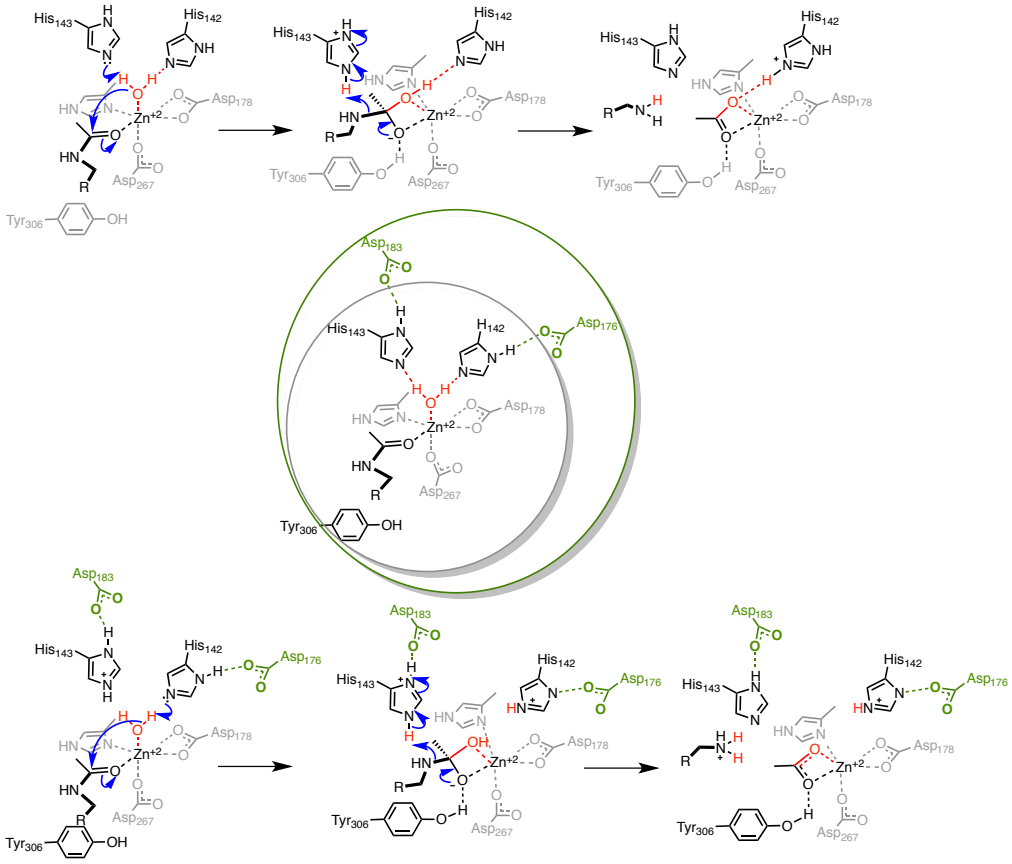


Figure 10. The active site matters: at the top, deacetylation in HDAC8 via the proposed proton shuttle mechanism is predicted using the smaller active site.[105] At the bottom, an extended active site model including two aspartate residues implicated the involvement of both histidines in the mechanism.[59, 106]

Sirtuin HDACs family: As observed for classical HDACs, MD calculations on proteins from the Sirtuin family show that broad conformational changes in the protein environment affect the epigenetic erasing process. Xu [108] studied allosteric stimulation of acetylation by Sir2 when bound to Sir4. 75 ns MD with the AMBER ff03 force field and SPC solvation revealed that this phenomenon is caused by favorable interactions between a long loop of Sir4 and the catalytic domain of Sir2. Lee [109] performed 5 ns MD with the GROMOS96 force field and SPC solvation that showed that the binding of the substrate and the NAD^+ to the apoform of SIRT2, a Sir2 homolog, directs a phenylalanine close to an active site histidine, therefore favoring the deacetylation. Chakrabarti [110] carried out 10 ns MD with the AMBER ff99SB

force field to conclude that the nicotinamide product of the acetylation inhibits the reaction by competing with the cofactor NAD^+.

The intrinsic details of deacetylation by proteins from the Sirtuin family have been computationally elucidated by Luo and Jiang [50, 111, 112] and Zhang [111, 112]. Luo and Jiang [50] suggested a S_N2 mechanism for the first step of the deacetylation by Sir2, after performing QM/MM calculations at the B3LYP//AMBER level of theory. Following 15 ns MD with the AMBER ff99 and SPC solvation, they reported that a glutamine and a histidine residue separately act as bases for the deprotonation step of the N-ribose and also help to the positioning of the cofactor NAD^+. The full reaction profile of the deacetylation reaction by Sir2 has been computationally determined by Zhang [111, 112] at the B3LYP//AMBER level of theory. They found that the first step, the nicotinamide cleavage, takes place through a concerted and highly dissociative displacement mechanism. The second step, highly dependent on a water molecule, is rate-determining and involves the formation of a bicyclic intermediate followed by its collapse. Through a qualitative analysis of conformation changes observed during MD-QM/MM simulations, Zhang [112] reported that the cofactor-binding loop is very dynamic and crucial for allowing water molecules to come close to the intermediates.

3.1.4 Uncovering histone lysine demethylation mechanisms and the role of the protein dynamics

Similarly to deacetylases, demethylases show structural flexibility that has been investigated by MD studies. Proteins from the LSD family, for instance, form complexes with other proteins and adopt different conformations that are proposed to affect demethylation. QM studies with these proteins have focused on understanding the multi-step reaction profile and the important contribution of individual residues upon the energetic profile, for instance intrinsic details of the hydrogen transfer step from lysine to the flavin cofactor, believed to be the rate-determining step in LSD proteins. Our recent work has focussed on the application of QM/MM calculations to JMJC proteins,

highlighting electrostatic contributions involved in O_2 binding which involve residues beyond the immediate active site surrounding the Fe-centre.[52]

JMJC family: Although proteins from the JMJC family share high active site similarity, differences in the protein residues in the vicinity of the active site significantly influence substrate binding and recognition.[52, 67] Gursoy [67] studied the Jumonji protein JMJD2A in complex with different lysine protonation states of the H3 histone by carrying out 18 ns MD simulations with the AMBER ff03 force field and explicit TIP3P solvation. It was found that both Coulombic and van der Waals interactions play an important role in binding of the di- and trimethylated H3 to JMJD2A, in particular a hydrogen bond between the side chain of the arginine residue and the backbone of the methylated lysine.

We have [52] performed 10 ns MD simulations with the AMBER ff99SB force filed and explicit TIP3P solvation and QM/MM calculations with B3LYP//AMBER to understand O_2 binding to the active site of the JMJD2A. We found that the computational description of residues not immediately attached to the non-heme Fe(II) centre are nonetheless essential to describe the energetics of the binding process (**Figure 11**). Small active site models may be useful to calculate the relative spin-state energetics of the bound complex. However, quantitative estimates of the binding affinity require a much larger model, which includes these more remote residues. Positively charged arginine and lysine residues play important electrostatic effects in the first step of the demethylation reaction because of the charge transfer that occurs upon O_2 binding. Based on these observations we suggested that mutation of these residues would impact the reaction energetics. By including the protein environment in our calculations, an additional 6.7 kcal mol^{-1} of electrostatic stabilisation is conferred to the overall binding free energy. We found that O_2 binding is slightly exergonic, even after inclusion of a quasi-harmonic term to account for the unfavourable entropy of association.

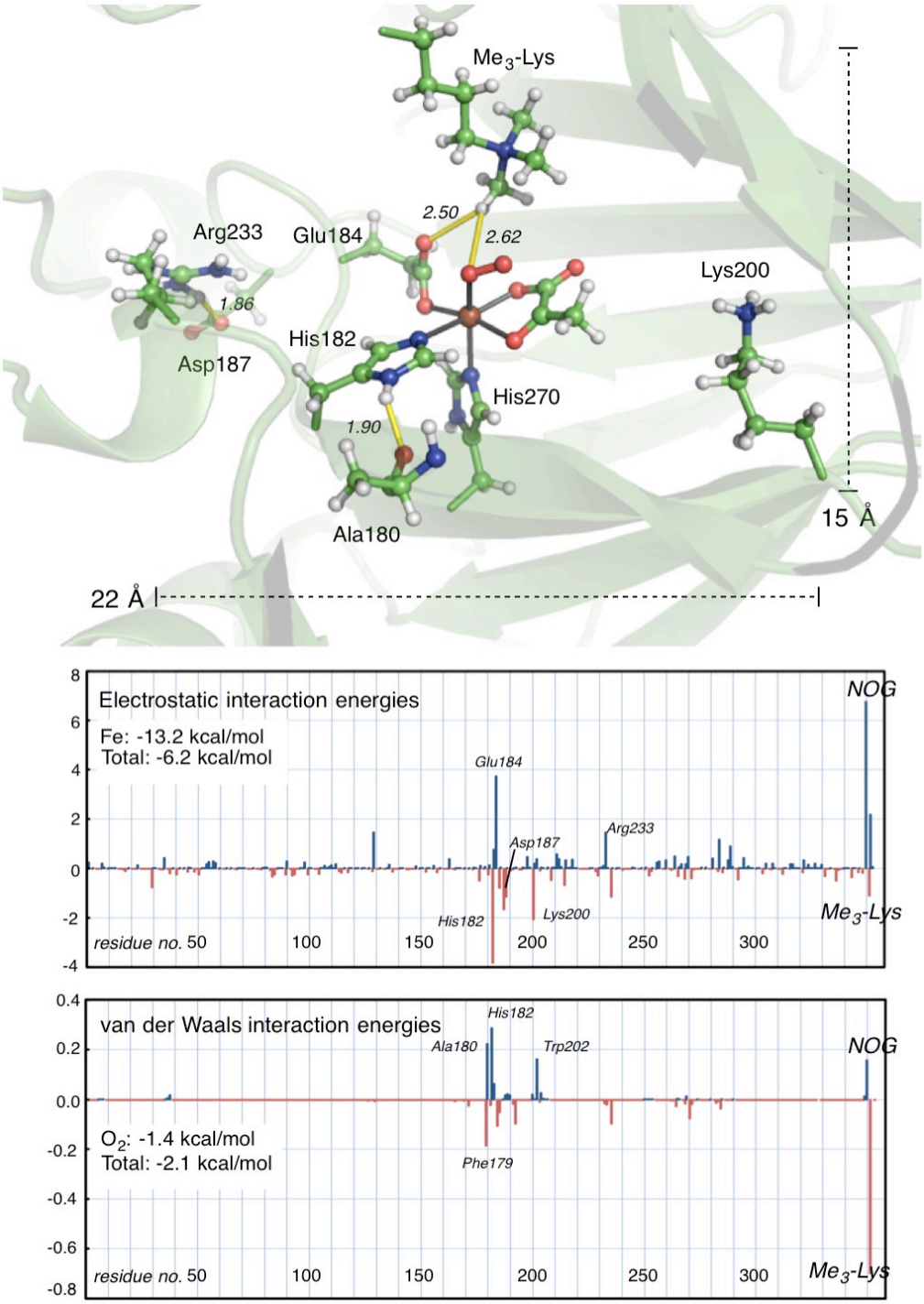


Figure 11. At the top, JMJD2A structure with O_2 bound.[52] Charged residues not directly bound to the active site, such as Lys200, Asp187 and Arg233 show significant electrostatic contributions for the total binding energy, as it can be seen in the EDA at the bottom. Taken from ref. [52]. *John Wiley & Sons*, Inc. Copyright @ 2015.

LSD family: A common feature observed in calculations with the erasing regulators is protein flexibility and its importance upon the chemical reaction. This has been highlighted in proteins belonging to the LSD family. Baron [113] studied LSD1 in complex with the REST corepressor 1 (CoREST), performing 50 ns MD with the 53A6 GROMOS force field and SPC solvation. A conserved active site lysine in the LSD family of proteins was found to have an important role for the diffusion of the O_2 into the active site, allowing the further steps of the demethylation reaction.

The LSD1/CoREST complex is highly dynamic and contains three domains: the amine oxidase (AO), the Tower and the Swirm domains. Baron [114] carried out microscale MD simulations with the 53A6 GROMOS force field and SPC solvation, finding that the unbound LSD1/CoREST presents a gate at the histone interaction site formed by a triad of lysine residues. They suggested that the complexation of LSD1/CoREST and histones affects the amino oxidase activity of LSD1. In another MD study also using the 53A6 GROMOS force field and SPC solvation, the same authors concluded that the histone binding changes the protein shape by an induced-fit mechanism, dramatically affecting the structural arrangements of the AO domain compared with the unbound LSD1/CoREST complex.[115]

QM/MM mechanistic studies on LSD proteins have revealed the most important residues involved in bond formation/cleavage associated with demethylation. Thiel [116] performed QM/MM calculations at the B3LYP//CHARMM22 level of theory to understand the LSD1 demethylation of histone H3. Three active site residues were found to be crucial for demethylation: a lysine acting as a base; a tyrosine helping the positioning of the reactants, and a tryptophan stabilizing the water bridge that plays a role during the proton transfer step. Jiang, Luo and Li [51] have also carried out 30 ns of MD simulation with the AMBER ff99SB force field and TIP3P solvation, showing the tyrosine has a stabilizing effect on the reaction. They performed QM/MM calculations at the B3LYP//AMBER level of theory to investigate the mechanistic profile. They found that the process involves hydride transfer

from the α–carbon of the lysine H3 to the flavin cofactor. This mechanism has also been studied by Truhlar [117] with electrostatically embedded multiconfiguration molecular mechanics (EE-MCMM). This approach approximates the gradient in a typical QM/MM calculation by an analytic potential based on a constant valence bond coupling term facilitating MD umbrella sampling. Their results show that this pathway happens via a concerted transfer of a hydrogen atom and an electron, with an activation free energy of 10 kcal mol^{-1}. Although formally equivalent to hydride transfer this description is preferred since the atomic charge on hydrogen remains positive throughout.

3.1.5 Recognition of Histone Acetylated Lysine

Computational mechanistic work on epigenetic reading regulators has been done mainly at the MM level, where the aim has been to identify residues that are important for substrate binding and specificity.

Pizzitutti [118] concluded that a proline residue on the bromodomain of Gcn5p (**Figure 12**) was essential for its ability to recognize acetylated lysine on H4 by affecting the movement and structure of the ZA loop, following 5-10 ns MD simulations with the CHARMM27 force field and TIP3P solvation. They also performed experimental mutation studies to confirm their hypothesis about the importance of the proline residue.

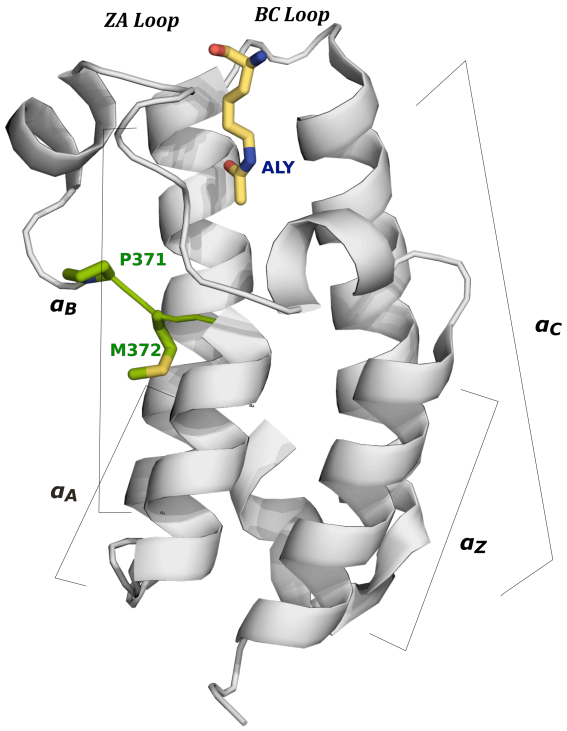


Figure 12. Gcn5p bromodomain. Shown here is the residue proline (P371, green), which although not in the binding site of the acetylated lysine is important for substrate recognition. This residue affects the dynamic and structure of the ZA loop (PDB ID: 1E6I).[118]

Caflisch [119, 120] and colleagues have used microscale MD with the CHARMM27 force field, and described the discovery of a “P-binding mode” in which the acetylated lysine forms a hydrogen bond to the backbone carbonyl of a proline located in the vacant lipophilic WPF shelf of the TBP-Associated Factor 1 (TAF1) bromodomain. During their simulations, they found the ligand reversibly interconverts with the well characterized “N-binding mode” in which the ligand is hydrogen bonded to a conserved asparagine residue. Interestingly the “P-binding mode” was observed in 50% of MD snapshots while the “N-binding mode” was found slightly less frequently in 40% of the snapshots. It is important to note that following Caflisch’s computational study, two independent studies experimentally identified a binding pose similar to the “P-binding mode” in other bromodomains.[121, 122] Some studies on the

acetyl lysine recognition process by bromodomains focused on the flexibility of the histone-bromodomain complexes. Caflisch [57] carried out 1 μs MD with the CHARMM22 force field. These simulations established a mechanism in which the promotion of opening of the bromodomain binding site is performed via a displacement of the side chains belonging to the conserved residues of the acetyl lysine-binding site. A tendency for the $–NH_2$ functional group of an asparagine to point towards bulk solvent was observed for significant amounts of time relative to the entire simulation time which was true for various receptors: TAF1L(2) (90% of the simulation time), TAF1(2) (81%), TAF1(1) (68%), Gcn5L2 (44%), ATAD2 (43%), BAZ2B (39%), and PB1(2) (38%). In the instances for TAF1L (2) and TAF1 bromodomains, the conformation was stabilized by hydrogen bonds mediated by water bridges with a serine. Five of the analysed bromodomains had a conserved tyrosine residue oriented away from the binding site for significant amounts of time in proteins BRDT(1) 100%, CREBBP (31%), BAZ2B (24%), SMARCA4 (21%), and BRD1(11%). Interestingly they report that this is only reflected in one crystal structure and one solution structure (PDB IDs: 2RFJ and 2DKW) out of more than 50 available crystal and solution structures. Flexibility has also been shown to allow the formation of an induced fit-pocket for small-molecule binding with CREBBP bromodomains.[123]

3.1.6 Recognition of Histone Methylated Lysine

The recognition process of methylated lysines usually involves a hydrophobic cage, formed by two to four aromatic residues (typically tryptophan, tyrosine and phenylalanine), which interact with the methylammonium ion through cation-π interactions [124, 125]. High states of methylation on lysines, such as di- and trimethylated lysines, tend to be recognized by readers containing such aromatic cages through a *surface-groove* mode, meaning that they are inserted in a wide and accessible binding site. Monomethylated lysines can also bind to these cages, but they usually require other negatively charged residues interacting with the non-methylated H-bond donors of the methyl lysine.[124, 126] Epigenetic readers achieve selectivity for monomethylated

substrates by having a binding site that contains a deep and size-sensitive protein cleft, known as the cavity-insertion recognition mode.

Two main themes emerge from computational work on these readers: i) understanding the flexibility of complexes formed between the reader regulators and the modified histones, ii) analysis of the importance of particular residues, mainly as part of hydrophobic cages, for the recognition and selectivity of different methylation states (e.g. mono-, di- or trimethylated lysines) of the histones. The influence of this cage in the recognition of different stated of methylation in lysine has been mainly studied by classical MD simulations, in a similar fashion to the bromodomain studies described above – most computational work has been performed for the chromodomains.

Zhou [127] performed 10 ns molecular dynamics with the AMBER ff03 force field and TIP3P solvation to evaluate the flexibility of the Drosophila heterochromatin-associated protein 1 (HP1) in the free and histone H3 bound states. Zhou found that the free complex is much more flexible than with bound methylated histones. It was also noticed that the trimethylated lysine interacts more favorably with the aromatic cage of HP1 than the monomethylated, thus providing a key feature for protein-protein recognition. These observations are in line with previous experimental studies showing the preference for H3K9Me_3-HP1 over lower methylation states.[128, 129] Wang [130] also evaluated stability of complexes formed between H3 and chromodomains, in this case the chromodomain *Drosophila melanogaster* protein polycomb. Firstly, they analysed the substrate selectivity through 4 ns molecular dynamics with the AMBER force field. They found from MM-GBSA binding free energy calculations that the most stable complex (the unmodified and different methylation states of K27 were considered computationally) is formed with the trimethylated H3, corroborating their experimental results. They also showed that H3K27Me_2 binds more weakly than H3K27Me_1 to polycomb, suggesting that dimethylated lysine is less likely to bind to the hydrophobic pocket of the chromodomain. Wang also performed an Electrostatic Decomposition Analysis to check which residues are more

important for the stability of the chromodomain-H3 complex. Aromatic tyrosine and tryptophan residues presented the greatest interactions with the peptide because they belong to a hydrophobic cage that has favourable interactions with the trimethylated lysine residue. In another work, Wang [131] showed that a hydrophobic cage, formed by a phenylalanine and a tryptophan, is important for the H3 binding to another chromodomain, the Chromobox Protein Homolog 6 Chromodomain (CBX6). 4 ns MD were carried out with the AMBER ff03 force field and TIP3P solvation. Still following the idea that hydrophobic cages are essential for substrate recognition, Wang [132] performed 2 ns MD with the AMBER ff03 force field and TIP3P solvation to conclude that hydrophobic residues tyrosine, phenylalanine and tryptophan, are fundamental for the stability of the chromodomain helicase DNA-binding Protein 1 with the methylated H3. Binding energy values for this complexation were slightly lower for the trimethylated histone than for the monomethylated histone.

Zhang studied the importance of charge independent effects as well as hydrophobic cage effects, that were previously explored by Wang, for H3 substrate recognition by the HP1 chromodomain with 5 ns MD using the AMBER ff99 force field.[133] They found cation-π interactions play an important role for the binding of the positively charged lysine compared to neutral substrates. On the other hand, hydrophobic cages are essential to distinguish the binding of the methylated H3 from the unmodified histone. Keskin [134] studied the tudor domains of the demethylase JMJD2A. They performed 25 ns MD with the AMBER ff03 force field to conclude that binding free energies are lower for the JMJD2A binding to H4 than to H3 and that binding is preferred to tri- than to dimethylated histones.

Dejaegere [135] analyzed the recognition of methylated H3 by PHD fingers and the most important energy contributions of individual residues. They carried out 10 ns MD using the CHARMM27 force field and TIP3P solvation. They highlighted that a conserved tryptophan residue is crucial, but alone is insufficient for histone recognition by PHD finger-containing proteins. It was

proposed that the first four residues of H3, known as the Linear Interaction Motif, and not only the modified lysine residue itself, are responsible for significant energetic contributions to the binding and recognition of the methylated substrate K4 to the PHD finger.

4. Conclusions and Outlook

Epigenetic processes are known to play a key role in a range of diseases, such as cancer and neurological disorders. Recent years have seen an increasing interest in developing small-molecule inhibitors targeting epigenetic regulators, with many of them entering the clinical-trial phase.

The development of more selective inhibitors requires an improved understanding of the chemical events triggering epigenetic processes. Computational simulations have been useful in helping to quantify and explain the atomic-scale mechanisms driving recognition and activity in these systems. They have also helped to establish the sequence of binding events and the consequences of protein dynamics upon binding and catalysis. Such knowledge has the potential to not only increase basic understanding, but also to provide a structural and mechanistic basis for the design of inhibitors, both in terms of efficiency and selectivity.

A particularly important aspect of the computational works covered in this review has been the consideration of the protein environment when modelling both binding and recognition of modified histone lysine residues by epigenetic proteins. They have demonstrated, for example, the role of loop-dynamics on histone modifying-reading proteins, where the protein can become active/inactive depending on how open/close they are for the substrate-binding step. Single residues can also act as a selective gate, by blocking channels that allow access of substrates and other water molecules important for the reaction. Finally, even when modelling the reaction itself, non-canonical residues have been shown to be crucial for positioning the substrate in an active binding mode. The choice of system partitioning in

QM/MM studies can also impact upon the resulting mechanistic picture provided by computation.

While there are still many open questions concerning recognition and activity in these systems, computational tools have provided significant insight into the chemical processes involved. Methodological aspects particularly relevant to consider in the continued design of small-molecule inhibitors targeting these systems are: i) long time MD simulations to understand the impact of loop-dynamic on recognition of methylated and acetylated histone lysines, ii) QM/MM studies, including the effects of dispersion in the active site and long-range electrostatic interactions from remote residues, for describing the chemical events involved, and iii) energy decomposition schemes to identify key residues contributing to recognition and catalysis.

With the maturation of the field of epigenetics and the advent of more powerful computational approaches, which could lead to increases in time scales and system sizes that can be studied, it is expected that atomistic simulations will become more important for understanding epigenetic-regulators and their interaction with specific inhibitors.

5. Acknowledgements

WAC is supported by a Science Without Borders (CAPES) scholarship. KK is supported by a World Bank Education Grant. FD acknowledges the Royal Society for a Newton International Fellowship. RSP acknowledges NVIDIA for computer hardware and the EPSRC UK National Service for Computational Chemistry Software (CHEM 870) for computational support. We acknowledge the expertise and constructive comments of reviewers in improving the final version of this Perspective.

6. References

[1] Brand, M., Measures, A.M., Wilson, B.G., Cortopassi, W.A., Alexander, R., Hoss, M., et al. Small Molecule Inhibitors of Bromodomain-Acetyl-lysine Interactions. Acs Chem Biol. 2015, 10, 22-39.

[2] Virani, S., Colacino, J.A., Kim, J.H., Rozek, L.S. Cancer epigenetics: a brief review. ILAR J. 2012, 53, 359-69.
[3] Phillips, D.M. The presence of acetyl groups of histones. Biochem J. 1963, 87, 258-63.
[4] Andreoli, F., Del Rio, A. Computer-aided Molecular Design of Compounds Targeting Histone Modifying Enzymes. Comput Struct Biotechnol J. 2015, 13, 358-65.
[5] Mersfelder, E.L., Parthun, M.R. The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure. Nucleic Acids Res. 2006, 34, 2653-62.
[6] Khoury, G.A., Baliban, R.C., Floudas, C.A. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. Sci Rep. 2011, 1.
[7] Kaminskas, E., Farrell, A., Abraham, S., Baird, A., Hsieh, L.S., Lee, S.L., et al. Approval summary: azacitidine for treatment of myelodysplastic syndrome subtypes. Clin Cancer Res. 2005, 11, 3604-8.
[8] Saba, H.I. Decitabine in the treatment of myelodysplastic syndromes. Ther Clin Risk Manag. 2007, 3, 807-17.
[9] VanderMolen, K.M., McCulloch, W., Pearce, C.J., Oberlies, N.H. Romidepsin (Istodax, NSC 630176, FR901228, FK228, depsipeptide): a natural product recently approved for cutaneous T-cell lymphoma. J Antibiot (Tokyo). 2011, 64, 525-31.
[10] Mann, B.S., Johnson, J.R., Cohen, M.H., Justice, R., Pazdur, R. FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. Oncologist. 2007, 12, 1247-52.
[11] Szyf, M. Prospects for the development of epigenetic drugs for CNS conditions. Nat Rev Drug Discov. 2015, 14, 461-74.
[12] Odore, E., Lokiec, F., Cvitkovic, E., Bekradda, M., Herait, P., Bourdel, F., et al. Phase I Population Pharmacokinetic Assessment of the Oral Bromodomain Inhibitor OTX015 in Patients with Haematologic Malignancies. Clin Pharmacokinet. 2015.
[13] Vellore, N.A., Baron, R. Epigenetic molecular recognition: a biomolecular modeling perspective. ChemMedChem. 2014, 9, 484-94.
[14] Smith, B.C., Denu, J.M. Chemical mechanisms of histone lysine and arginine modifications. Biochim Biophys Acta. 2009, 1789, 45-57.
[15] Falkenberg, K.J., Johnstone, R.W. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. Nat Rev Drug Discov. 2014, 13, 673-91.
[16] Hojfeldt, J.W., Agger, K., Helin, K. Histone lysine demethylases as targets for anticancer therapy. Nat Rev Drug Discov. 2013, 12, 917-30.
[17] Filippakopoulos, P., Knapp, S. Targeting bromodomains: epigenetic readers of lysine acetylation. Nat Rev Drug Discov. 2014, 13, 337-56.
[18] McAllister, T.E., England, K.S., Hopkinson, R.J., Brennan, P.E., Kawamura, A., Schofield, C.J. Recent Progress in Histone Demethylase Inhibitors. J Med Chem. 2015.
[19] Arrowsmith, C.H., Bountra, C., Fish, P.V., Lee, K., Schapira, M. Epigenetic protein families: a new frontier for drug discovery. Nat Rev Drug Discov. 2012, 11, 384-400.

[20] Yoo, J., Medina-Franco, J.L. Inhibitors of DNA methyltransferases: insights from computational studies. Curr Med Chem. 2012, 19, 3475-87.
[21] Aldeghi, M., Heifetz, A., Bodkin, M.J., Knapp, S., Biggin, P.C. Accurate calculation of the absolute free energy of binding for drug molecules. Chem Sci. 2016, 7, 207-18.
[22] Yan, C., Xiu, Z., Li, X., Li, S., Hao, C., Teng, H. Comparative molecular dynamics simulations of histone deacetylase-like protein: binding modes and free energy analysis to hydroxamic acid inhibitors. Proteins. 2008, 73, 134-49.
[23] Karaman, B., Sippl, W. Docking and binding free energy calculations of sirtuin inhibitors. Eur J Med Chem. 2015, 93, 584-98.
[24] Wang, M., Mok, M.W., Harper, H., Lee, W.H., Min, J., Knapp, S., et al. Structural genomics of histone tail recognition. Bioinformatics. 2010, 26, 2629-30.
[25] Martin, C., Zhang, Y. The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol. 2005, 6, 838-49.
[26] Graff, J., Tsai, L.H. Histone acetylation: molecular mnemonics on the chromatin. Nat Rev Neurosci. 2013, 14, 97-111.
[27] Marino-Ramirez, L., Kann, M.G., Shoemaker, B.A., Landsman, D. Histone structure and nucleosome stability. Expert Rev Proteomics. 2005, 2, 719-29.
[28] Chin, H.G., Patnaik, D., Esteve, P.O., Jacobsen, S.E., Pradhan, S. Catalytic properties and kinetic mechanism of human recombinant Lys-9 histone H3 methyltransferase SUV39H1: participation of the chromodomain in enzymatic catalysis. Biochemistry. 2006, 45, 3272-84.
[29] Patnaik, D., Chin, H.G., Esteve, P.O., Benner, J., Jacobsen, S.E., Pradhan, S. Substrate specificity and kinetic mechanism of mammalian G9a histone H3 methyltransferase. J Biol Chem. 2004, 279, 53248-58.
[30] Tanner, K.G., Langer, M.R., Kim, Y., Denu, J.M. Kinetic mechanism of the histone acetyltransferase GCN5 from yeast. J Biol Chem. 2000, 275, 22048-55.
[31] Berndsen, C.E., Albaugh, B.N., Tan, S., Denu, J.M. Catalytic mechanism of a MYST family histone acetyltransferase. Biochemistry. 2007, 46, 623-9.
[32] Pandey, R., Muller, A., Napoli, C.A., Selinger, D.A., Pikaard, C.S., Richards, E.J., et al. Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes. Nucleic Acids Res. 2002, 30, 5036-55.
[33] Marmorstein, R. Structure of histone acetyltransferases. J Mol Biol. 2001, 311, 433-44.
[34] Yan, Y., Harper, S., Speicher, D.W., Marmorstein, R. The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate. Nat Struct Biol. 2002, 9, 862-9.
[35] Yan, Y., Barlev, N.A., Haley, R.H., Berger, S.L., Marmorstein, R. Crystal structure of yeast Esa1 suggests a unified mechanism for catalysis and substrate binding by histone acetyltransferases. Mol Cell. 2000, 6, 1195-205.
[36] Liu, X., Wang, L., Zhao, K., Thompson, P.R., Hwang, Y., Marmorstein, R., et al. The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. Nature. 2008, 451, 846-50.
[37] Hamamoto, R., Saloura, V., Nakamura, Y. Critical roles of non-histone protein lysine methylation in human tumorigenesis. Nat Rev Cancer. 2015, 15, 110-24.
[38] Bai, Q., Shen, Y., Yao, X., Wang, F., Du, Y., Wang, Q., et al. Modeling a new water channel that allows SET9 to dimethylate p53. PLoS One. 2011, 6, e19856.

[39] Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell. 2004, 119, 941-53.
[40] Stavropoulos, P., Blobel, G., Hoelz, A. Crystal structure and mechanism of human lysine-specific demethylase-1. Nat Struct Mol Biol. 2006, 13, 626-32.
[41] Kooistra, S.M., Helin, K. Molecular mechanisms and potential functions of histone demethylases. Nat Rev Mol Cell Biol. 2012, 13, 297-311.
[42] Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., et al. Histone demethylation by a family of JmjC domain-containing proteins. Nature. 2006, 439, 811-6.
[43] Ng, S.S., Kavanagh, K.L., McDonough, M.A., Butler, D., Pilka, E.S., Lienard, B.M., et al. Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. Nature. 2007, 448, 87-91.
[44] Langley, G.W., Brinko, A., Munzel, M., Walport, L.J., Schofield, C.J., Hopkinson, R.J. Analysis of JmjC Demethylase-Catalyzed Demethylation Using Geometrically-Constrained Lysine Analogues. Acs Chem Biol. 2015.
[45] Walport, L.J., Hopkinson, R.J., Schofield, C.J. Mechanisms of human histone and nucleic acid demethylases. Curr Opin Chem Biol. 2012, 16, 525-34.
[46] Marks, P.A., Xu, W.S. Histone deacetylase inhibitors: Potential in cancer therapy. J Cell Biochem. 2009, 107, 600-8.
[47] Yang, X.J., Seto, E. HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. Oncogene. 2007, 26, 5310-8.
[48] Lombardi, P.M., Cole, K.E., Dowling, D.P., Christianson, D.W. Structure, mechanism, and inhibition of histone deacetylases and related metalloenzymes. Curr Opin Struct Biol. 2011, 21, 735-43.
[49] Sauve, A.A. Sirtuins. Biochim. Biophys Acta. 2010, 1804, 1565-6.
[50] Liang, Z., Shi, T., Ouyang, S., Li, H., Yu, K., Zhu, W., et al. Investigation of the catalytic mechanism of Sir2 enzyme with QM/MM approach: S_N1 vs S_N2? J Phys Chem B. 2010, 114, 11927-33.
[51] Kong, X., Ouyang, S., Liang, Z., Lu, J., Chen, L., Shen, B., et al. Catalytic mechanism investigation of lysine-specific demethylase 1 (LSD1): a computational study. PLoS One. 2011, 6, e25444.
[52] Cortopassi, W.A., Simion, R., Honsby, C.E., Franca, T.C., Paton, R.S. Dioxygen Binding in the Active Site of Histone Demethylase JMJD2A and the Role of the Protein Environment. Chem. Eur. J. 2015.
[53] Denis, G.V. Bromodomain coactivators in cancer, obesity, type 2 diabetes, and inflammation. Discov Med. 2010, 10, 489-99.
[54] Muller, S., Filippakopoulos, P., Knapp, S. Bromodomains as therapeutic targets. Expert Rev Mol Med. 2011, 13, e29.
[55] Filippakopoulos, P., Knapp, S. The bromodomain interaction module. FEBS Lett. 2012, 586, 2692-704.
[56] Eissenberg, J.C. Structural biology of the chromodomain: form and function. Gene. 2012, 496, 69-78.
[57] Steiner, S., Magno, A., Huang, D., Caflisch, A. Does bromodomain flexibility influence histone recognition? FEBS Lett. 2013, 587, 2158-63.
[58] Vreven, T., Morokuma, K., Farkas, O., Schlegel, H.B., Frisch, M.J. Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints. J Comput Chem. 2003, 24, 760-9.

[59] Chen, K., Zhang, X., Wu, Y.D., Wiest, O. Inhibition and mechanism of HDAC8 revisited. J Am Chem Soc. 2014, 136, 11636-43.
[60] Duangkamol Gleeson, M.P.G. Application of QM/MM and QM methods to investigate histone deacetylase 8. Medicinal Chemical Communications. 2015, 6, 477-85.
[61] Jiang, J., Lu, J., Lu, D., Liang, Z., Li, L., Ouyang, S., et al. Investigation of the acetylation mechanism by GCN5 histone acetyltransferase. PLoS One. 2012, 7, e36660.
[62] Zhang, X., Ouyang, S., Kong, X., Liang, Z., Lu, J., Zhu, K., et al. Catalytic mechanism of histone acetyltransferase p300: from the proton transfer to acetylation reaction. J Phys Chem B. 2014, 118, 2009-19.
[63] Grimme, S., Antony, J., Ehrlich, S., Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys. 2010, 132, 154104.
[64] Hirao, H. The effects of protein environment and dispersion on the formation of ferric-superoxide species in myo-inositol oxygenase (MIOX): a combined ONIOM(DFT:MM) and energy decomposition analysis. J Phys Chem B. 2011, 115, 11278-85.
[65] Lonsdale, R., Harvey, J.N., Mulholland, A.J. Effects of Dispersion in Density Functional Based Quantum Mechanical/Molecular Mechanical Calculations on Cytochrome P450 Catalyzed Reactions. J Chem Theory Comput. 2012, 8, 4637-45.
[66] Carvalho, A.T., Gouveia, M.L., Raju Kanna, C., Warmlander, S.K., Platts, J., Kamerlin, S.C. Theoretical modelling of epigenetically modified DNA sequences. F1000Res. 2015, 4, 52.
[67] Ulucan, O., Keskin, O., Erman, B., Gursoy, A. A comparative molecular dynamics study of methylation state specificity of JMJD2A. PLoS One. 2011, 6, e24664.
[68] Genshaft, A., Moser, J.A., D'Antonio, E.L., Bowman, C.M., Christianson, D.W. Energetically unfavorable amide conformations for N6-acetyllysine side chains in refined protein structures. Proteins. 2013, 81, 1051-7.
[69] Tanner, K.G., Trievel, R.C., Kuo, M.H., Howard, R.M., Berger, S.L., Allis, C.D., et al. Catalytic mechanism and function of invariant glutamic acid 173 from the histone acetyltransferase GCN5 transcriptional coactivator. J Biol Chem. 1999, 274, 18157-60.
[70] Sternglanz, R., Schindelin, H. Structure and mechanism of action of the histone acetyltransferase Gcn5 and similarity to other N-acetyltransferases. Proc Natl Acad Sci U S A. 1999, 96, 8807-8.
[71] Jedrzejewski, R.P., Kazmierkiewicz, R. Structure of Patt1 human proapoptotic histone acetyltransferase. J Mol Model. 2013, 19, 5533-8.
[72] Di Fenza, A., Rocchia, W., Tozzini, V. Complexes of HIV-1 integrase with HAT proteins: multiscale models, dynamics, and hypotheses on allosteric sites of inhibition. Proteins. 2009, 76, 946-58.
[73] Quy, V.C., Pantano, S., Rossetti, G., Giacca, M., Carloni, P. HIV-1 Tat Binding to PCAF Bromodomain: Structural Determinants from Computational Methods. Biology (Basel). 2012, 1, 277-96.
[74] Mujtaba, S., He, Y., Zeng, L., Farooq, A., Carlson, J.E., Ott, M., et al. Structural basis of lysine-acetylated HIV-1 Tat recognition by PCAF bromodomain. Mol Cell. 2002, 9, 575-86.

[75] Terreni, M., Valentini, P., Liverani, V., Gutierrez, M.I., Di Primio, C., Di Fenza, A., et al. GCN5-dependent acetylation of HIV-1 integrase enhances viral integration. Retrovirology. 2010, 7, 18.
[76] Roth, S.Y., Denu, J.M., Allis, C.D. Histone acetyltransferases. Annu Rev Biochem. 2001, 70, 81-120.
[77] Cyr, A.R., Domann, F.E. The redox basis of epigenetic modifications: from mechanisms to functional consequences. Antioxid Redox Signal. 2011, 15, 551-89.
[78] Zhang, X., Yang, Z., Khan, S.I., Horton, J.R., Tamaru, H., Selker, E.U., et al. Structural basis for the product specificity of histone lysine methyltransferases. Mol Cell. 2003, 12, 177-85.
[79] Wu, H., Mathioudakis, N., Diagouraga, B., Dong, A., Dombrovski, L., Baudat, F., et al. Molecular basis for the regulation of the H3K4 methyltransferase activity of PRDM9. Cell Rep. 2013, 5, 13-20.
[80] Dillon, S.C., Zhang, X., Trievel, R.C., Cheng, X. The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol. 2005, 6, 227.
[81] Zhang, X., Wen, H., Shi, X. Lysine methylation: beyond histones. Acta Biochim Biophys Sin (Shanghai). 2012, 44, 14-27.
[82] Kaniskan, H.U., Jin, J. Chemical probes of histone lysine methyltransferases. Acs Chem Biol. 2015, 10, 40-50.
[83] Chuikov, S., Kurash, J.K., Wilson, J.R., Xiao, B., Justin, N., Ivanov, G.S., et al. Regulation of p53 activity through lysine methylation. Nature. 2004, 432, 353-60.
[84] Kwon, T., Chang, J.H., Kwak, E., Lee, C.W., Joachimiak, A., Kim, Y.C., et al. Mechanism of histone lysine methyl transfer revealed by the structure of SET7/9-AdoMet. EMBO J. 2003, 22, 292-303.
[85] Huang, J., Perez-Burgos, L., Placek, B.J., Sengupta, R., Richter, M., Dorsey, J.A., et al. Repression of p53 activity by Smyd2-mediated methylation. Nature. 2006, 444, 629-32.
[86] Zhang, X., Bruice, T.C. Histone lysine methyltransferase SET7/9: formation of a water channel precedes each methyl transfer. Biochemistry. 2007, 46, 14838-44.
[87] Zhang, X., Bruice, T.C. Mechanism of product specificity of AdoMet methylation catalyzed by lysine methyltransferases: transcriptional factor p53 methylation by histone lysine methyltransferase SET7/9. Biochemistry. 2008, 47, 2743-8.
[88] Cheng, X., Collins, R.E., Zhang, X. Structural and sequence motifs of protein (histone) methylation enzymes. Annu Rev Biophys Biomol Struct. 2005, 34, 267-94.
[89] Min, J., Feng, Q., Li, Z., Zhang, Y., Xu, R.M. Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. Cell. 2003, 112, 711-23.
[90] Qiao, Q., Li, Y., Chen, Z., Wang, M., Reinberg, D., Xu, R.M. The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. J Biol Chem. 2011, 286, 8361-8.
[91] Hu, P., Zhang, Y. Catalytic mechanism and product specificity of the histone lysine methyltransferase SET7/9: an ab initio QM/MM-FE study with multiple initial structures. J Am Chem Soc. 2006, 128, 1272-8.

[92] Guo, H.B., Guo, H. Mechanism of histone methylation catalyzed by protein lysine methyltransferase SET7/9 and origin of product specificity. Proc Natl Acad Sci U S A. 2007, 104, 8797-802.
[93] Wang, S., Hu, P., Zhang, Y. Ab initio quantum mechanical/molecular mechanical molecular dynamics simulation of enzyme catalysis: the case of histone lysine methyltransferase SET7/9. J Phys Chem B. 2007, 111, 3758-64.
[94] Hu, P., Wang, S., Zhang, Y. How do SET-domain protein lysine methyltransferases achieve the methylation state specificity? Revisited by Ab initio QM/MM molecular dynamics simulations. J Am Chem Soc. 2008, 130, 3806-13.
[95] Trievel, R.C., Beach, B.M., Dirk, L.M., Houtz, R.L., Hurley, J.H. Structure and catalytic mechanism of a SET domain protein methyltransferase. Cell. 2002, 111, 91-103.
[96] Chu, Y., Yao, J., Guo, H. QM/MM MD and free energy simulations of G9a-like protein (GLP) and its mutants: understanding the factors that determine the product specificity. PLoS One. 2012, 7, e37674.
[97] Xu, Q., Chu, Y.Z., Guo, H.B., Smith, J.C., Guo, H. Energy triplets for writing epigenetic marks: insights from QM/MM free-energy simulations of protein lysine methyltransferases. Chemistry. 2009, 15, 12596-9.
[98] Chu, Y., Sun, L., Zhong, S. How Y357F, Y276F mutants affect the methylation activity of PRDM9: QM/MM MD and free energy simulations. J Mol Model. 2015, 21, 125.
[99] Haider, S., Joseph, C.G., Neidle, S., Fierke, C.A., Fuchter, M.J. On the function of the internal cavity of histone deacetylase protein 8: R37 is a crucial residue for catalysis. Bioorg Med Chem Lett. 2011, 21, 2129-32.
[100] Kalyaanamoorthy, S., Chen, Y.P. Ligand release mechanisms and channels in histone deacetylases. J Comput Chem. 2013, 34, 2270-83.
[101] Kalyaanamoorthy, S., Chen, Y.P. Exploring inhibitor release pathways in histone deacetylases using random acceleration molecular dynamics simulations. J Chem Inf Model. 2012, 52, 589-603.
[102] Thangapandian, S., John, S., Lee, Y., Arulalapperumal, V., Lee, K.W. Molecular modeling study on tunnel behavior in different histone deacetylase isoforms. PLoS One. 2012, 7, e49327.
[103] Arrar, M., Turnham, R., Pierce, L., de Oliveira, C.A., McCammon, J.A. Structural insight into the separate roles of inositol tetraphosphate and deacetylase-activating domain in activation of histone deacetylase 3. Protein Sci. 2013, 22, 83-92.
[104] Kunze, M.B., Wright, D.W., Werbeck, N.D., Kirkpatrick, J., Coveney, P.V., Hansen, D.F. Loop interactions and dynamics tune the enzymatic activity of the human histone deacetylase 8. J Am Chem Soc. 2013, 135, 17862-8.
[105] Wu, R., Wang, S., Zhou, N., Cao, Z., Zhang, Y. A proton-shuttle reaction mechanism for histone deacetylase 8 and the catalytic role of metal ions. J Am Chem Soc. 2010, 132, 9471-9.
[106] Vanommeslaeghe, K., De Proft, F., Loverix, S., Tourwe, D., Geerlings, P. Theoretical study revealing the functioning of a novel combination of catalytic motifs in histone deacetylase. Bioorg Med Chem. 2005, 13, 3987-92.
[107] Finnin, M.S., Donigian, J.R., Cohen, A., Richon, V.M., Rifkind, R.A., Marks, P.A., et al. Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. Nature. 1999, 401, 188-93.

[108] Hsu, H.C., Wang, C.L., Wang, M., Yang, N., Chen, Z., Sternglanz, R., et al. Structural basis for allosteric stimulation of Sir2 activity by Sir4 binding. Genes Dev. 2013, 27, 64-73.
[109] Sakkiah, S., Arooj, M., Cao, G.P., Lee, K.W. Insight the C-site pocket conformational changes responsible for sirtuin 2 activity using molecular dynamics simulations. PLoS One. 2013, 8, e59278.
[110] Guan, X., Lin, P., Knoll, E., Chakrabarti, R. Mechanism of inhibition of the human sirtuin enzyme SIRT3 by nicotinamide: computational and experimental studies. PLoS One. 2014, 9, e107729.
[111] Hu, P., Wang, S., Zhang, Y. Highly dissociative and concerted mechanism for the nicotinamide cleavage reaction in Sir2Tm enzyme suggested by ab initio QM/MM molecular dynamics simulations. J Am Chem Soc. 2008, 130, 16721-8.
[112] Shi, Y., Zhou, Y., Wang, S., Zhang, Y. Sirtuin Deacetylation Mechanism and Catalytic Role of the Dynamic Cofactor Binding Loop. J Phys Chem Lett. 2013, 4, 491-5.
[113] Baron, R., Binda, C., Tortorici, M., McCammon, J.A., Mattevi, A. Molecular mimicry and ligand recognition in binding and catalysis by the histone demethylase LSD1-CoREST complex. Structure. 2011, 19, 212-20.
[114] Baron, R., Vellore, N.A. LSD1/CoREST is an allosteric nanoscale clamp regulated by H3-histone-tail molecular recognition. Proc Natl Acad Sci U S A. 2012, 109, 12509-14.
[115] Vellore, N.A., Baron, R. Molecular dynamics simulations indicate an induced-fit mechanism for LSD1/CoREST-H3-histone molecular recognition. BMC Biophys. 2013, 6, 15.
[116] Karasulu, B., Patil, M., Thiel, W. Amine oxidation mediated by lysine-specific demethylase 1: quantum mechanics/molecular mechanics insights into mechanism and role of lysine 661. J Am Chem Soc. 2013, 135, 13400-13.
[117] Yu, T., Higashi, M., Cembran, A., Gao, J., Truhlar, D.G. Concerted hydrogen atom and electron transfer mechanism for catalysis by lysine-specific demethylase. J Phys Chem B. 2013, 117, 8422-9.
[118] Pizzitutti, F., Giansanti, A., Ballario, P., Ornaghi, P., Torreri, P., Ciccotti, G., et al. The role of loop ZA and Pro371 in the function of yeast Gcn5p bromodomain revealed through molecular dynamics and experiment. J Mol Recognit. 2006, 19, 1-9.
[119] Spiliotopoulos, D., Caflisch, A. Molecular Dynamics Simulations of Bromodomains Reveal Binding-Site Flexibility and Multiple Binding Modes of the Natural Ligand Acetyl-Lysine. Isr. J. Chem. 2014, 54, 1084-92.
[120] Magno, A., Steiner, S., Caflisch, A. Mechanism and Kinetics of Acetyl-Lysine Binding to Bromodomains. J Chem Theory Comput. 2013, 9, 4225-32.
[121] Ferguson, F.M., Fedorov, O., Chaikuad, A., Philpott, M., Muniz, J.R., Felletar, I., et al. Targeting low-druggability bromodomains: fragment based screening and inhibitor design against the BAZ2B bromodomain. J. Med. Chem. 2013, 56, 10183-7.
[122] Lucas, X., Wohlwend, D., Hugle, M., Schmidtkunz, K., Gerhardt, S., Schule, R., et al. 4-Acyl pyrroles: mimicking acetylated lysines in histone code reading. Angew Chem Int Ed. 2013, 52, 14055-9.
[123] Rooney, T.P., Filippakopoulos, P., Fedorov, O., Picaud, S., Cortopassi, W.A., Hay, D.A., et al. A series of potent CREBBP bromodomain ligands reveals an

induced-fit pocket stabilized by a cation-pi interaction. Angew Chem Int Ed. 2014, 53, 6126-30.
[124] Carey, N., Royal Society of Chemistry (Great Britain). Epigenetics for drug discovery. Cambridge, UK, Royal Society of Chemistry, 2016.
[125] Gao, C., Herold, J.M., Kireev, D. Assessment of free energy predictors for ligand binding to a methyllysine histone code reader. J Comput Chem. 2012, 33, 659-65.
[126] Kamps, J.J., Huang, J., Poater, J., Xu, C., Pieters, B.J., Dong, A., et al. Chemical basis for the recognition of trimethyllysine by epigenetic reader proteins. Nat Commun. 2015, 6, 8911.
[127] Yanke Jiang, J.Z., Min Zeng, Na Zhang, Qingsen Yu. Dynamics simulation on the flexibility and backbone motions of HP1 chromodomain bound to free and lysine 9-methylated histone H3 tails. Int J Quantum Chem. 2009, 109, 1135–47.
[128] Jacobs, S.A., Khorasanizadeh, S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. Science. 2002, 295, 2080-3.
[129] Fischle, W., Tseng, B.S., Dormann, H.L., Ueberheide, B.M., Garcia, B.A., Shabanowitz, J., et al. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. Nature. 2005, 438, 1116-22.
[130] Stein, R.S., Li, N., He, W., Komives, E., Wang, W. Recognition of methylated peptides by Drosophila melanogaster polycomb chromodomain. J Proteome Res. 2013, 12, 1467-77.
[131] Li, N., Stein, R.S., He, W., Komives, E., Wang, W. Identification of methyllysine peptides binding to chromobox protein homolog 6 chromodomain in the human proteome. Mol Cell Proteomics. 2013, 12, 2750-60.
[132] Stein, R.S., Wang, W. The recognition specificity of the CHD1 chromodomain with modified histone H3 peptides. J Mol Biol. 2011, 406, 527-41.
[133] Lu, Z., Lai, J., Zhang, Y. Importance of charge independent effects in readout of the trimethyllysine mark by HP1 chromodomain. J Am Chem Soc. 2009, 131, 14928-31.
[134] Ozboyaci, M., Gursoy, A., Erman, B., Keskin, O. Molecular recognition of H3/H4 histone tails by the tudor domains of JMJD2A: a comparative molecular dynamics simulations study. PLoS One. 2011, 6, e14765.
[135] Grauffel, C., Stote, R.H., Dejaegere, A. Molecular dynamics for computational proteomics of methylated histone H3. Biochim Biophys Acta. 2015, 1850, 1026-40.