Deacetylase inhibitors & Histone inheritance

Zwinderman, Martijn R. H.

DOI:
10.33612/diss.167867692

Chapter 1

Ik hoef u niet te vertellen, dat de opgerolde dubbelspiraal DNA met de informatie voor een mens, die hermetische caduceus die in de kern van elk van zijn honderdduizend miljard cellen zit, niet meer dan een honderdduizendste gram weegt, maar dat zij, uitgerold, ongeveer even lang is als de mens zelf: het aantal mogelijke volgorden op moleculair niveau is dus een onverdragelijk getal. Geschreven in de drie-letterwoorden van het vier-letteralfabet wordt een mens bepaald door een genetisch verhaal, waarmee het equivalent van vijfhonderd bijbels gevuld kan worden.

Harry Mulisch, De ontdekking van de hemel (1992)

I don't have to tell you that the coiled double DNA helix containing the information on a human individual, that Hermetic caduceus within the nucleus of each of the individual's hundred thousand billion cells, weighs no more than one hundred thousandth of a gram but, when extended, is approximately the same length as the individual himself, so that the number of possible sequences at the molecular level is vast. If written in the three-letter words of the four-letter alphabet, a human being is determined by a genetic narrative long enough to fill the equivalent of five hundred Bibles.

Harry Mulisch, The Discovery of Heaven (1992)

Introduction

Two years before Harry Mulisch completed his The Discovery of Heaven, and one month before I was born, on October the first 1990 to be precise, the Human Genome Project was launched with the goal of determining the sequence of base pairs that make up human DNA or in other words, to decipher "the language in which God created life" as Bill Clinton announced upon completion of the project 13 years later [1]. Its completion, not coincidentally, coincided with the 50th anniversary of Watson and Crick's conclusion that DNA has a three-dimensional double-helix structure [2]. With more than 2,800 researchers that collaborated around the globe it remains the world's largest biological project to date. The Human Genome Project sequenced approximately 90% of the human genetic narrative, which was met with the belief that we had made a paradigm shift towards a better understanding of human disease. However, since then, we have come to the realization that life is more complex than originally thought. It's no longer sufficient to think of DNA as the Holy Grail to understand disease. In fact, one could argue that simply understanding DNA has opened a Pandora's Box and that the real work has only just begun as we start to delve further into disease mechanisms and pathobiology.

Actually, well before the start of the Human Genome Project we knew that life is incredibly more complex than can be discerned from knowing the sequence of DNA. Already in 1957 Crick presented a set of basic assumptions in an one-hour lecture that later became the central dogma in biology [3, 4]. At the basis he stated that "once information has got into a protein [indirectly from the genetic code] it can't get out again. Information here means the sequence of the amino acid residues, or other sequences related to it", meaning that the information in a DNA sequence cannot be rewritten from information in a protein [5]. Moreover, it was clear to Crick that information can flow from DNA to RNA and from RNA to protein. He argued that the latter flow of information would require an adaptor molecule, which he speculated "could

consist of perhaps a single chain of RNA. Each [RNA] adaptor molecule containing, say, a di- or trinucleotide would each be joined to its own amino acid by a special enzyme. These molecules would then diffuse to the microsomal particles and attach to the proper place on the basis of the RNA by base-pairing [5]." The existence of such RNA adaptor molecules was confirmed one year later and was called transfer RNA (tRNA) [6]. Each tRNA molecule translates a specific combination of three RNA nucleotides (Harry Mulisch's three-letter words, known as a codon in biology) into one of twenty canonical amino acids that make up a protein. This completes the flow of information from DNA to protein in two steps involving transcription of DNA into messenger RNA (mRNA) and translation of mRNA into amino acids using tRNA. The central dogma in biology illustrates in its simplest way the complexity of life and, besides, is a truly inspiring example of the predictive power of conceptual thinking.

The central dogma further tells us that DNA is just one form of information and thereby just one level at which things can go wrong and lead to disease. Of course, being at the top of the information cascade the genome has an extremely important function and mistakes in the DNA sequence can lead to serious diseases. However, many diseases cannot be clearly linked to mere mistakes in the DNA sequence but are rather the result of information transfer going wrong at a different level, for instance, at the level of transcription. Zooming in on the genome, we see that DNA is a long biopolymer of nucleotides that is wrapped around a repeating ball of eight proteins called histones in the form of chromatin, like a string wrapped around beads. The string can be coiled tightly or lightly around the beads to distinguish between "open" and "closed" chromatin and serves as a conceptual model between actively transcribed and repressed genomic regions, respectively. Variability in DNA packaging around histones, tightly or lightly, is mostly the result of differences in posttranslational modification (PTM) of histones [7], which is a form of information transfer that was not explicitly covered by Crick's central dogma. It involves editing of the histones with

molecular groups by appropriate enzymes. Numerous histone residues can be modified in a variety of ways, including ubiquitination, phosphorylation, acetylation and methylation, each having their own specific effect, but invariably influencing the state of the chromatin, allowing either more or less transcription of DNA into mRNA [8]. The pattern of histone modifications along the genome thereby acts as an additional code on top of (epi) the genetic code, like bookmarks in the genetic narrative, structuring and annotating the genome. This way the histone code provides an important explanation of how cells with the same DNA are able to differentiate into distinct cell lines.

To give you an idea of the complexity of the histone code, one histone (isoform H3) contains 19 lysine residues that can either be acetylated with one acetyl-group or methylated with up to three methyl-groups. Assuming that the modifications are independent of each other, this allows for the unbearable number of 5^{19} or 19 trillion different acetylation and methylation patterns. Not to mention including the other possible modifications nor the other histone isoforms. Furthermore, this code can be written, read and erased by a specific set of enzymes and is therefore much more dynamic than the genetic code. It becomes increasingly clear that the histone code is an important requisite for cell-specific gene expression and that changes in the histone code are associated with cancer and ageing [9]. In our attempts to unravel the epigenetic layer of information we have come to the realization that we are more than just our DNA than we ever dared to imagine.

Protein lysine acetylation and deacetylation

In view of the seemingly limitless number of histone modification patterns and my limited time as a researcher, I have focused on histone acetylation and the set of enzymes that erase the acetyl mark: the histone deacetylases (HDACs). Histone acetylation only occurs on lysine residues in histones. Lysines are positively charged at physiological pH and are present in the N-terminal DNA-binding tail of histones. Acetylation is thought to facilitate transcriptional activation either by neutralization of the ionic interaction between the lysine-containing histone tail and DNA or by forming a binding

site for chromatin remodeling proteins [10]. Moreover, besides histones, principally any protein with a lysine can undergo a dynamic process of acetylation and deacetylation at some point during its lifetime [11]. One study mapped over 3,600 of such acetylation sites on approximately 1,750 proteins [12]. This so-called acetylome is under the control of acetyltransferases and deacetylases, of which many isoforms exist. To date, 18 deacetylases are known, which are grouped into two families based on their dependency on either zinc or nicotinamide adenine dinucleotide as cofactor and are further divided into four different classes based on sequence similarity [13]. Deacetylases that contain zinc as cofactor are still mostly called HDACs, named after their firstly discovered target. However, since HDACs are also able to deacetylate non-histone proteins they are more aptly referred to as deacetylases, thus without the prefix histone. Yet, in keeping with current literature I will mostly refer to them as HDACs throughout this thesis. The functionally opposing group of proteins that show acetyltransferase activity is even larger and more diverse [14].

Epidrugs

The enzymes that install and remove PTMs on histones can be considered to be the principal executors of epigenetic mechanisms. Being highly dynamic in nature, PTMs are not only regulated in a cell-specific but also in a time-specific manner, in turn turning many different genes on or off over the course of a cell cycle. The crucial consequence of this discovery is that gene expression can directly be controlled at any point in time by drugs that target the enzymes that carry out PTMs. To refer to their actions on epigenetic processes such drugs are often termed “epidrugs” [15]. “Epidrug” is, however, not an all-encompassing name that only refers to an epigenetic component. It becomes increasingly clear that also non-histone proteins are targeted by enzymes that were initially discovered to act on histones [16]. Therefore, the actions on non-histone proteins need always be taken into account when evaluating the effects of epidrugs.

The potential of deacetylases as drug targets

Protein acetylation influences important protein properties. Next to altering DNA-protein interactions, protein-protein interactions, subcellular localization and even transcriptional activity can all be tuned by the dynamic process of acetylation [17–19]. Acetylation of the transcription factor p53 for example destabilizes the interaction with its negative regulator Mdm2, thereby activating p53, and ultimately leading to an increase in the expression of apoptosis and growth arrest-inducing genes [20]. In a similar fashion, acetylation of the p65 subunit of nuclear factor-kappa B (NF-κB), an important inflammatory transcription factor [21], influences its subcellular location and transcriptional activity [22]. Since acetylation is a posttranslational modification with fundamental importance for the function of key proteins, it is not surprising that acetylation patterns are distorted in various diseases. In cancer, p53 is acetylated to a lesser extent, allowing cells to grow rapidly [23]. In ulcerative colitis and other inflammatory diseases, the acetylation status of NF-κB is changed, resulting in an increased expression of cytokines [18, 19]. Consequently, small molecule inhibitors of HDACs (HDACi), which increase acetylation by repressing deacetylation, are in over 100 clinical trials for cancer and may also enter clinical trials for inflammatory diseases [26]. So far, four HDACi have been FDA-approved and these are known to relatively selectively induce apoptosis in a number of tumor cell types [27]. As many of the underlying processes of cancer are also important in inflammation, HDACi may be equally useful in chronic inflammatory diseases [28, 29]. Importantly, their anti-inflammatory effects generally occur at 10- to 100-fold lower concentrations than needed for their cytotoxic effects on cancer cells [30].

The currently FDA-approved HDACi are pan-HDAC inhibitors, which cause an overall increase in acetylation. Many side effects are expected to originate from this non-selectivity. To improve HDACi in oncology it is therefore important to develop inhibitors that target individual deacetylase isoforms [31]. It is thought that isoenzyme selectivity will be even more crucial in non-oncological applications such as inflammation, however, the individual

contributions of the various deacetylase isoforms in specific disease models need to be understood to enable their exploitation as therapeutic targets.

Additionally, apart from their enzymatic activity, HDACs also have roles in the formation of protein-protein complexes. It is important to be aware of this when comparing conditional knock-out studies to studies using HDAC inhibitors [32]. This is exemplified by the HDACs 4, 5, 7 and 9 for which their deacetylase activity might not be crucial for their function [33]. HDAC4, for instance, is unable to efficiently deacetylate proteins on its own [33]. Instead, this deacetylase binds HDAC3 and the thus resulting HDAC4/HDAC3 complex has deacetylase activity [32]. Additionally, these HDACs shuttle between the nucleus and the cytoplasm and thus their subcellular localization, rather than their expression or activity *per se*, controls their actions.

Thesis outline

Part one of this thesis aims to define the possible utility of HDACi in the treatment of diseases with an inflammatory component. In **chapter two** the particular roles of zinc-dependent lysine deacetylases in models of asthma, chronic obstructive pulmonary disease (COPD) and lung cancer are explained. Additionally, the role of protein and DNA methylation in these diseases is described. Subsequently, **chapter three** provides an overview of current structure-activity relationships of HDACi selectively targeting HDAC1, 2 and 3 complexes in disease. Finally, **chapter four** demonstrates that optimization of the inhibitory profile among HDACs 1, 2 and 3 is important for application of these inhibitors in diseases that require either immune inhibition, such as in inflammatory diseases, or immune activation, such as in immuno-oncology.

Part two aims to unravel patterns of histone inheritance and starts in **chapter five** with an overview of bioorthogonal chemistries that have enabled the dissemination of complex biological processes through metabolic labeling strategies. In **chapter six** a newly developed metabolic labeling procedure that tracks new histones during DNA replication is used to investigate whether

new histones distribute evenly on duplicated DNA strands. Additionally, the effect of replication stress on the deposition of new histones during replication is presented. Finally, in chapter seven a summary of the main findings of this thesis is provided together with an outlook towards areas of future research.

References

1. Blighe K, DeDionisio L, Christie KA, et al (2018) Gene editing in the context of an increasingly complex genome. BMC Genomics 19:595

2. Watson JD, Crick FHC (1953) Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. Nature 171:737–738

3. Crick FHC (1958) On protein synthesis. Symp Soc Exp Biol 12:138–163

4. Crick FHC (1970) Central dogma of molecular biology. Nature 227:561–563

5. Cobb M (2017) 60 years ago, Francis Crick changed the logic of biology. PLoS Biol. 15:

6. Hoagland MB, Stephenson ML, Scott JF, Hecht LI, Zamecnik PC (1958) A soluble ribonucleic acid intermediate in protein synthesis. J Biol Chem 231:241–257

7. Li B, Carey M, Workman JL (2007) The Role of Chromatin during Transcription. Cell 128:707–719

8. Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. Cell Res 21:381–395

9. Venkatesh S, Workman JL (2015) Histone exchange, chromatin structure and the regulation of transcription. Nat Rev Mol Cell Biol 16:178–189

10. Becker PB, Workman JL (2013) Nucleosome remodeling and epigenetics. Cold Spring Harb. Perspect. Biol. 5:

11. Drazic A, Myklebust LM, Ree R, Arnesen T (2016) The world of protein acetylation. Biochim Biophys Acta - Proteins Proteomics 1864:1372–1401

12. Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, Olsen J V, Mann M (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science 325:834–840

13. De Ruijter AJM, Van Gennip AH, Caron HN, Kemp S, Van Kuilenburg ABP (2003) Histone deacetylases (HDACs): Characterization of the classical HDAC family. Biochem J 370:737–749

14. Yang XJ, Seto E (2007) HATs and HDACs: From structure, function and regulation to novel strategies for therapy and prevention. Oncogene 26:5310–5318

15. Mai A, Altucci L (2009) Epi-drugs to fight cancer: From chemistry to cancer treatment, the road ahead. Int J Biochem Cell Biol 41:199–213

16. Altucci L, Rots MG (2016) Epigenetic drugs: from chemistry via biology to medicine and back. Clin Epigenetics 8:1–3

17. Verdin E, Ott M (2015) 50 years of protein acetylation: From gene regulation to epigenetics, metabolism and beyond. Nat Rev Mol Cell Biol 16:258–264

18. Choudhary C, Weinert BT, Nishida Y, Verdin E, Mann M (2014) The growing landscape of lysine acetylation links metabolism and cell signalling. Nat Rev Mol Cell Biol 15:536–550

19. Spange S, Wagner T, Heinzel T, Krämer OH (2009) Acetylation of non-histone proteins modulates cellular signalling at multiple levels. Int J Biochem Cell Biol 41:185–198

20. Tang Y, Zhao W, Chen Y, Zhao Y, Gu W (2008) Acetylation Is Indispensable for p53 Activation. Cell 133:612–626

21. Lawrence T (2009) The nuclear factor NF-kappaB pathway in inflammation. Cold Spring Harb Perspect Biol 1:1–10

22. Huang B, Yang X-D, Lamb A, Chen L-F (2010) Posttranslational modifications of NF-kappaB: another layer of regulation for NF-kappaB signaling pathway. Cell Signal 22:1282–90

23. Reed SM, Quelle DE (2014) P53 acetylation: Regulation and consequences. Cancers (Basel) 7:30–69

24. Glauben R, Siegmund B (2011) Inhibition of histone deacetylases in inflammatory bowel diseases. Mol Med 17:426–433

25. Ito K, Charron CE, Adcock IM (2007) Impact of protein acetylation in inflammatory lung diseases. Pharmacol Ther 116:249–265

26. Mottamal M, Zheng S, Huang TL, Wang G (2015) Histone deacetylase inhibitors in clinical studies as templates for new anticancer agents. Molecules 20:3898–3941

27. Bolden JE, Shi W, Jankowski K, Kan CY, Cluse L, Martin BP, MacKenzie KL, Smyth GK, Johnstone RW (2013) HDAC inhibitors induce tumor-cell-selective pro-apoptotic transcriptional responses. Cell Death Dis 4:e-519

28. Coussens LM, Werb Z (2002) Inflammation and cancer. Nature 420:860–867

29. Shakespear MR, Halili MA, Irvine KM, Fairlie DP, Sweet MJ (2011) Histone deacetylases as regulators of inflammation and immunity. Trends Immunol

32:335–343

30. Dinarello CA, Fossati G, Mascagni P (2011) Histone Deacetylase Inhibitors for Treating a Spectrum of Diseases Not Related to Cancer. Mol Med 17:333–352

31. Cantley MD, Haynes DR (2013) Epigenetic regulation of inflammation: Progressing from broad acting histone deacetylase (HDAC) inhibitors to targeting specific HDACs. Inflammopharmacology 21:301–307

32. Fischle W, Dequiedt F, Hendzel MJ, Guenther MG, Lazar MA, Voelter W, Verdin E (2002) Enzymatic activity associated with class II HDACs is dependent on a multiprotein complex containing HDAC3 and SMRT/N-CoR. Mol Cell 9:45–57

33. Lahm A, Paolini C, Pallaoro M, et al (2007) Unraveling the hidden catalytic activity of vertebrate class IIa histone deacetylases. Proc Natl Acad Sci 104:17335–17340